# Bernoulli News

Newsletter of the Bernoulli Society For Mathematical Statistics and Probability

Vol. 24 (1), May 2017 | Published twice per year by the Bernoulli Society | ISSN 1360–6727


## CONTENTS




**Editor**

MIGUEL DE CARVALHO
School of Mathematics
THE UNIVERSITY *of* EDINBURGH
EDINBURGH, UK

**Contact**

*miguel.decarvalho@ed.ac.uk*


† *Bernoulli News is the official newsletter of the Bernoulli Society, publishing news, calendars of events, and opinion pieces of interest to Bernoulli Society members, as well as to the Mathematical Statistics and Probability community at large. The views and opinions expressed in editorials and opinion pieces do not necessarily reflect the official views of the Bernoulli Society, unless explicitly stated, and their publication in Bernoulli News in no way implies their endorsement by the Bernoulli Society. Consequently, the Bernoulli Society does not bear any responsibility for the views expressed in such pieces.*


## A VIEW FROM THE PRESIDENT

Dear Members of the Bernoulli Society,

As we all seem to agree, the role and image of statistics has changed dramatically. Still, it takes ones breath when realizing the huge challenges ahead.

Statistics has not always been considered as being very necessary. In 1848 the Dutch Ministry of Home Affairs established an office of statistics. And then, thirty years later minister Kappeyne van de Coppelo abolishes the "superfluous" office. The office was quite rightly put back in place in 1899, as "Centraal Bureau voor de Statistiek" (CBS). Statistics at the CBS has evolved from "simple" counting to an art requiring a broad range of competences. Of course counting remains important. For example the CBS reports in February 2017 that almost 1 out of 4 people entitled to vote in the Netherlands is over the age of 65. But clearly, knowing this generates questions. What is the influence of this on the outcome of the elections? This calls for more data. Demographic data are combined with survey data and nowadays also with data from other sources, in part to release the "survey pressure" that firms and individuals are facing. Getting sensible information from the the resulting complicated data heap requires a careful and ingenious analysis.

Also for example the US Census Bureau "reuses data from other agencies to cut the cost of data collection and to reduce the burden on people who respond to our censuses and surveys." (`https://www.census.gov/about/what.html`). In other words, governmental offices take part in the Big Data Story and hence need more and more highly educated experts. Already the question which data are to be combined, reused or ignored is eminently nontrivial.

Big Data Analytics have also entered the the courtroom and judges are to be assisted by experts in statistical analysis. Courts, industry, electronic commerce, its all about collecting, addressing, storing and analyzing data, and most importantly, about a joint effort to deal with data.

The former President of the Swiss Statistical Society Diego Kuonen says when interviewed by Alison Oliverat (see: `https:/goo.gl/MNIv1x`):

"Data science is moving very fast and statisticians are not the fast guys but we have the chance to jump on the train and to put it in the right direction. Collaborating is very important as there is a lot of denial going on. Statistics is key to data collection and some only see it in terms of its use when it comes to confirmatory or explanatory analysis using p-values for example. There are plenty of opportunities ahead. This really is a golden age for statistics."

Collaboration is needed, also within the various areas of statistics. Due the complexity of Data Science, cross fertilization is crucial. The borders between e.g. official, mathematical, computational, industrial statistics, statistics and education, statistics and privacy, statistics and ethics, are even less defined than before. The 61st World Statistics Congress ISI2017 in Marrakech has this wonderful mix of areas. It provides a great opportunity to learn from each other and start joint projects (and do not forget to attend the General Assembly there).


… *Continued on p. 1*



***Deadline for the next issue:*** **30 September, 2017**
Send contributions to: *miguel.decarvalho@ed.ac.uk*

# A View from the President (continued from front cover)

Let me quote Willem van Zwet who tells about his travel by train to the ISI 1961 in Paris (see: R. J. Beran and N. I. Fisher "An Evening Spent with Bill van Zwet," *Statistical Science*, 2009, 24, 87–115):

> "In our compartment, there was a gentleman who told us he was also going to attend the ISI Session, so I thought we'd come across a kindred soul. He went on to explain that mathematics was a lot of nonsense. All you needed was to collect data and they would speak for themselves. This was my first acquaintance with official statistics and it came as a bit of a shock. Much later I found out that most government statisticians have a broader outlook."

Today, we ARE kindred souls.

Times are changing.

In the same interview, and after mentioning the first five successful European Meetings of Statisticians, Willem van Zwet notes that something went wrong during the sixth version at Hannover and that it looked like the European Meetings had died. Now look what happened (and thank Willem)! Have a look in this issue for the 31st European Meeting of Statisticians in Helsinki! At the same time, have a look at 39th Conference on Stochastic Processes and their Applications in Moscow! Then there is the delicate decision which one you will attend. Or use your modern technology and attend them both! Well, my time as Bernoulli President is almost over. Someone fantastic is taking over the stick at the Bernoulli General Assembly in Marrakech. Susan, I wish you a wonderful Presidency and look forward to your term!

*Sara van de Geer*
*President of the Bernoulli Society*
*Zurich*

## Editorial

Corina Constantinescu is our new e-Briefs editor. Welcome! Many thanks to Leonardo T. Rolla for his wonderful job as e-Briefs editor over the last two years. This is the first issue of *Bernoulli News* typeset in LaTeX.

# News from the Bernoulli Society

## Nominations for European Regional Committee

Eight nominations have been received to fill the eight vacancies on the European Regional Committee (ERC) of the Bernoulli Society. These are:

- Andreas Basse-O'Connor, Aarhus University, Denmark.
- Geurt Jongbloed, Delft University, The Netherlands.
- Peter Kevei, University of Szeged, Hungary.
- Tatyana Krivobokova, Göttingen, Germany.
- Marloes Maathuis, ETH Zürich, Switzerland.
- Davy Paindaveine, ULB Brussels, Belgium.
- Laura Sangalli, Politecnico di Milano, Italy.
- Ulrike Schneider, Vienna University of Technology, Austria.

According to the ERC statutes, any group of at least five European members of the Bernoulli Society is entitled to nominate further candidates who have declared themselves willing to serve on the Committee. Additional nominations should be sent by email to r.samworth@statslab.cam.ac.uk, and would then force an election. If no further nominations are received by Friday 2 June, then the eight candidates listed above will be declared elected.

*Richard Samworth*
*Cambridge*

# Awards and Prizes

## Doeblin Prize 2016


Allan Sly (recipient of Doeblin Prize 2016) and Sara van de Geer.

During the 2016 Bernoulli World Congress in Toronto, Allan Sly was awarded the Doeblin Prize. The prize is sponsored by Springer. Sara van de Geer as Bernoulli Society President added the following to accompany the awarding of the prize:

"This prize is awarded to a single individual for outstanding research in the field of probability, and who is at the beginning of his or her mathematical career. The Committee for Conferences on Stochastic Processes has elected you as the prize winner.

You are a very broad young probabilist, with great taste for fundamental problems at the intersection between probability theory, statistics, statistical physics and theoretical computer science. You have worked on topics such as Glauber dynamics for the Ising model, random k-satisfiability, interacting particle systems, the connection between computational complexity and phase transitions. Each of these topics has witnesed fundamental progress due to your efforts.

You are invited to submit to the journal Probability Theory and Related Fields a paper for publication as the Wolfgang Doeblin Prize Article, and will also be invited to present the Doeblin Prize Lecture at Conference on Stochastic Processes and their Applications 2017.

Then, I hereby hand over the certificate, as well as an official letter of the BS.

Congratulations!"

*Sara van de Geer*
*President of the Bernoulli Society*
*Zurich*

# New Executive Members in the Bernoulli Society

## Chair of the European Regional Committee: Niels Richard Hansen

Niels Richard Hansen obtained his PhD in 2004 from University of Copenhagen. He has been employed at the Department of Mathematical Sciences, University of Copenhagen since 2005, and he is currently Professor in Computational Statistics. From 2010 to 2011 he was Visiting Scholar in the Biostatistics Group, University of Berkeley. His main research topics are multivariate and high-dimensional dynamic models; model selection and penalized estimation methods; causal dynamic models; machine learning; and R software development. His interests include theory and computational methodology as well as biological applications. Niels was chair of the Danish Society for Theoretical Statistics from 2008 to 2010, and he has been a member of the European Regional Committee from 2012 to 2016. He became an elected member of ISI in 2016. He is currently the editor of *Scandinavian Journal of Statistics*.

### *Niels's View on the European Regional Committee*

It is a great pleasure to take over the role as chair of the European Regional Committee (ERC) from Richard Samworth, and I will use this opportunity to thank him for serving as chair and for the good collaboration I have had with him in ERC. According to the statutes the object of ERC is "to promote European cooperation in the sciences of probability theory and statistics and their applications," and the main way to achieve this is to initiate the organization of conferences and meetings. Three major events under the auspices of ERC are the European Meetings of Statisticians (in Helsinki, July 24–28, 2017), the European Young Statisticians Meeting (in Uppsala, August 14–18, 2017), and the Seminaires Europeens de Statistiques (SemStat) series of short courses. The SemStat series has thanks to Ernst Wit been revived with the course Statistical Network Science at Eurandom, March 7–10, 2017.

The ERC is completely dependent on the many local organizers and program committee members, who invest their time in organizing the meetings and courses. It is therefore important for me to express how highly appreciated these investments are by the ERC. Without your efforts there would be no meetings! I look forward to seeing you as well as all the other participants at one or more of the meetings.

## Bernoulli Youth Representative: Parthanil Roy

Parthanil Roy is an associate professor in the Theoretical Statistics and Mathematics Division of Indian Statistical Institute. He obtained his PhD (2007) from Cornell University on Stable Random Fields. He was a postdoctoral fellow (2007–08) at the RiskLab, ETH Zurich and an assistant professor (2008–11) at Michigan State University before joining Indian Statistical Institute, where he was promoted to an associate professor position in Nov. 2015. Parthanil's research focuses on the interplay between probability theory and ergodic theory in the context of heavy tails, stable processes, long range dependence and branching random walks. He received Microsoft Young Faculty Award (2012) and was selected as an Associate of Indian Academy of Sciences (2012–15). Presently, he is an associate editor of *Sankhya, Ser.* A, and serving in the scientific committee of Heavy Tails and Long Range Dependence Conference to be held in Paris during June 20–22, 2017.

### *Parthanil's View on the Role of Bernoulli Youth Representative*

It is an honour to take over the position of Bernoulli Youth Representative from Corina Constantinescu, who has done a wonderful job as a bridge between young researchers and Bernoulli Society. In a time when every academic society and organization is facing the challenge of recruiting and retaining young academicians, the role of a youth representative becomes extremely important. We need to attract the attention of the younger generation of probabilists and statisticians, and convince them how important it is to get involved in an organization like Bernoulli Society from both professional and social viewpoint. If anyone has any suggestion on how to attain this goal, please feel free to e-mail me at parthanilroy@isibang.ac.in.

# Articles and Letters

## On Bayesian Measures of Uncertainty in Large or Infinite-Dimensional Models

Judith Rousseau, Université Paris Dauphine, France
rousseau@ceremade.dauphine.fr

Communicated by Sara van de Geer

This letter summarizes key ideas from the Ethel Newbold prize lecture. Recent questions and approaches in Bayesian nonparametrics are discussed, along with Bernstein–von Mises results for nonparametric and semi-parametric models, as well as frequentist coverage of nonparametric credible balls.

### *§1. Bayesian Nonparametrics: Approach and Questions*

Bayesian nonparametrics is a rapidly growing domain of research, with applications in all sorts of fields, such as biostatistics, physics, economics social science, computation biology etc. It makes this area of research highly exciting whether one is interested in modeling, implementation or theory.

In Bayesian statistics one considers a sampling model for the observations, say $X \in \mathcal{X} \sim P_\theta$, with a parameter $\theta \in \Theta$, together with a probability distribution $\Pi$ on the parameter space $\Theta$, called the prior probability. This prior probability $\Pi$ allows to incorporate in the inference all sorts of prior information which is available on the parameter $\theta$ before observing $X$. It also allows to model the (prior) uncertainty on the parameter. The inference is then based on the posterior distribution which is the conditional probability distribution of the parameter $\theta$ given the observation $X$. Using this posterior distribution one can construct point estimators, test procedures and measures of uncertainty such as credible regions or estimation of losses (or risks).

In large dimensional models, it is impossible to construct a prior in a fully informative way, i.e. based solely on prior information. In the same time, the more complex the model, the more difficult it is to apprehend the impact of the prior distribution on the inference. As an illustration, consider a parameter $\theta$ which is a function belonging to $L^2[0,1]$, so that $\Theta = L^2[0,1]$ and a prior probability on $L^2[0,1]$ to be a zero mean Gaussian process on $[0,1]$. The covariance kernel of the Gaussian process is a crucial component of the prior modeling; in Figure 1 we show draws from two Gaussian processes: One from the Brownian motion on $[0,1]$ and the other from a Gaussian process with exponential kernel $K(u,v) = e^{-(u-v)^2}$.

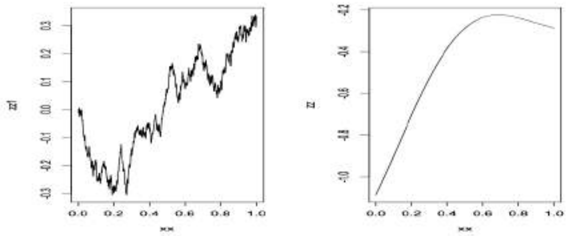


Figure 1: Gaussian processes: Brownian motion (left) and process based on exponential kernel (right).

These are typical draws from the prior and obviously they are very different but would they lead to very different inference procedures when combined with a likelihood (sampling model)?

Moreover, one of the most interesting features of statistics is that it does more than provide point estimates or predictors by constructing reliable measures of uncertainty, without which inference makes little sense. Bayesian approaches, by defining a probability on the parameter space, naturally allow for the construction and definition of such measures of uncertainty on $\theta$ or on anything related to $X$ and $\theta$. But these raise a number of questions:

1 In the setup of Figure 1, would the inference remain different between the two types of covariance kernels when the number $n$ of observations or the information in the sample $X$ increases to infinity? More generally, given a particular structure of a family of prior distributions, can we detect features of the prior model which will remain influential asymptotically?

2 Are the measures of uncertainty robust to the choice of the prior and to which extent?

3 Which implicit assumptions are made on the parameter by a given prior?

Studying the asymptotic behaviour of the posterior distribution is a way to answer at least partially these questions. Indeed it leads to a better understanding of the impact of the prior distribution since any feature which remains influential asymptotically is likely to have very strong influence for finite samples and the way they act on the posterior distribution asymptotically is a good indicator on the way they act at finite $n$. There is now a large literature on asympotic properties of the posterior distribution with a particular focus on the so-called posterior consistency or posterior concentration rate, following the seminal papers of Barron et al. (1999) and Ghosal et al. (2000) which paved the way to this line of research. However these results only partly inform on the properties of the Bayesian measures of uncertainty and a more refined analysis is needed to understand their behaviour. Indeed if $\epsilon_n$ is the posterior concentration rate at $\theta_0 \in \Theta$ for some metric $d(\cdot, \cdot)$, then

$$\Pi\left(d(\theta_0, \theta) \leq \epsilon_n | X\right) = 1 + o_{P_{\theta_0}}(1),$$

and under quite general conditions a Bayesian $\alpha$ credible ball in the form $\{\theta; d(\hat{\theta}, \theta) \leq q_\alpha(X)\}$ will have radius bounded by $O(\epsilon_n)$ but we do not know anything about its frequentist coverage. Hence posterior concentration rates are informative about the size of credible regions but are not informative about their frequentist coverage, see for instance Hoffman et al. (2015).

### §2. Are Credible Regions also Confidence Regions?

An $\alpha$- credible region is any measurable subset $C$ of the parameter space $\Theta$ such that the posterior probability of $C$, $\Pi(C|X)$, is equal (or at least larger than) to $1 - \alpha$. There exist infinitely many ways to construct credible regions, see for instance Robert (2001), however commonly used credible regions are either: balls (or bands or ellipsoids) centered at the posterior mean or highest posterior density regions. In regular finite dimensional parametric models these regions are also asymptotic confidence regions, i.e. they verify for any compact subset $K$ of $\Theta$,

$$\inf_{\theta \in K} P_\theta\left(\theta \in C\right) = 1 - \alpha + o(1). \qquad (1)$$

This implies that asymptotically the prior distribution has no influence and all lead to the same inference on the parameter : either for point estimation or measures of uncertainty. Unfortunately (1) does not extend easily to large or infinite-dimensional models. Interestingly, thanks to Fubini's theorem, any $\alpha$ credible region is on average an $\alpha$ confidence region, i.e. it satisfies

$$\int_\Theta P_\theta\left(\theta \in C\right) d\Pi(\theta) = 1 - \alpha. \qquad (2)$$

However, in large or infinite-dimensional models (2) conveys little information on the set of parameter values for which (1) is satisfied or at least for which the frequentist coverage $P_\theta\left(\theta \in C\right)$ is bounded from below by some reasonable constant and is therefore of little use to answer the second question of Section 1. In the last 5 years or so, some advances have been obtained in this direction for nonparametric and semiparametric models, but much if not most remains to be done. In the following sections I will summarize the results which have been obtained.

#### §2.1. Bernstein–von Mises Results for Nonparametric and Semi-parametric Models

In parametric models, the main tool which leads to (1), is the renown Bernstein–von Mises theorem which states that the posterior distribution is asymptotically Gaussian with mean the maximum likelihood estimator $\hat{\theta}$ (or any efficient estimator) and variance the inverse of the Fisher information matrix, in other words the frequentist variance of $\hat{\theta}$. This duality between the asymptotic form of the posterior distribution and the frequentist distribution of the posterior mean implies that, to first order, Bayesian and frequentist (likelihood) inference are equivalent so that Bayesian and frequentist measures of uncertainty coincide asymptotically. Hence a way to derive (1) is to extend the Bernstein–von Mises theorem to infinite-dimensional models. The first results in this direction were negative, see Cox (1993) and Freedman (1999). Recently, weaker versions of the Bernstein–von Mises have been obtained in infinite-dimensional models, following essentially two types of approaches. The first type corresponds to deriving the Bernstein–von Mises theorem for finite dimensional functionals of the parameter: $\psi : \Theta \to \mathbb{R}^d$ for some $d \geq 1$. For instance in partially linear models the regression function of a response variable $x$ on a covariate vector $z$, $f(z)$, is decomposed into a linear part $z_1^t\beta$, $\beta \in \mathbb{R}^d$ and a non linear nuisance parameter part $h(z_2)$, $z = (z_1, z_2)$ and the parameter of interest is $\beta$. In survival analysis a commonly used model is the Cox model where the hazard rate has the form $h(x)e^{z^t\beta}$, with $h$ the unknown baseline hazard rate and the parameter of interest is typically $\beta$, so that $\psi(\theta) = \beta$ with $\theta = (\beta, h)$. There are also other examples where $\theta$ does not naturally write as $(\psi, h)$, for instance $\psi(\theta) = \|\theta\|^2$ or $\psi(\theta) =< a, \theta >$ for some (function) $a$. In both types of problems there are many instances where the parameter of interest $\psi(\theta)$ can be estimated at the (parametric) rate $1/\sqrt{n}$; see for instance van der Vaart (1998) for a general theory on semiparametric efficient estimation. The Bayesian counterpart of the theory of semi-paramatric efficiency can be obtained by studying the asymptotic posterior distribution of $\sqrt{n}(\psi(\theta) - \hat{\psi})$ where $\hat{\psi}$ is an efficient estimator of $\psi(\theta)$. When it is asymptotically normal with zero mean and variance $V_0$ where $V_0$ is the asymptotic frequentist variance of $\hat{\psi}$, the posteriori distribution on $\psi(\theta)$ is said to satisfy a Bernstein–von Mises theorem (BvM) and (1) holds for credible regions $C$ associated to the parameter $\psi(\theta)$.

In Castillo (2010) and Bickel and Kleijn (2012) sufficient conditions are proposed to derive BvM in models in the form $\theta = (\psi, h)$. In Rivoirard and Rousseau (2012) and in Castillo and Rousseau (2015) sufficient conditions to BvM are provided for linear functional of the density and for smooth functionals of the parameter in general models respectively.

Taking a slightly different approach, Castillo and Nickl (2013) and Castillo and Nickl (2014) have derive general conditions for a BvM to hold on the whole parameter $\theta$ but under a weaker norm (than the weak convergence of the posterior distribution). This is an elegant approach which allows in particular to derive BvM for some smooth functionals of the parameter, as in Castillo and Rousseau (2015).

In all these papers, although some non trivial models are studied the families of prior models that have been considered is rather limited. One of the reasons is that the conditions which need to be verified are quite involved and require a good understanding of the nature of Kullback–Leibler type neighbourhoods of the true parameter $\theta_0$. When the prior model induces a non regular geometry, which is for instance the case of most nonparametric mixture models these conditions are very difficult to check and proving a Bersntein–von Mises theorem for such models is still an open problem.

### *§2.2. Frequentist Coverage of Nonparametric Credible Balls*

It might be argued that to derive a statement such as (1), a Bernstein–von Mises result is not necessary and that it might prove more efficient to study directly the frequentist coverage of $\alpha$-credible balls in the form $C = \{\theta, d(\theta, \hat{\theta}) \leq q_\alpha\}$, where $q_\alpha$ is defined by $\Pi(C|X) = 1 - \alpha$ and $\hat{\theta}$ is a well chosen estimator such as the posterior mean. This is the route which has been considered for instance in Szabó et al. (2015).

For such problems, when $\Theta$ is infinite dimensional there are (common) cases where it is known that it is not possible to construct confidence regions $C$ which have optimal size and uniformly correct frequentist coverage. These cases appear typically when $\Theta$ is a collection of embedded sets $\Theta_\alpha$ where $\alpha$ is some index of regularity or sparsity of the parameter. As seen previously the size of credible ball is bounded by the posterior concentration rate, and in many cases this rate is optimal so that Bayesian credible balls are often (if the prior is correctly calibrated) optimal in size, hence they cannot enjoy correct uniform frequentist coverage. In Szabó et al. (2015), the authors determine a subset of $\Theta = \ell_2$ over which (1) is satisfied, which they call the polish tail set of parameters for the Gaussian white noise model with hierarchical Gaussian priors. This idea, which is similar in spirit to what is done also in the frequentist literature, has then been extended last year to some high-dimensional sparse problems such as the Gaussian many means model, or the high-dimensional linear regression model, see Nurushev and Belitser (2016); Belitser and Ghosal (2016); van der Pas et al. (2017). It has also been generalized to some generic nonparametric models in Rousseau and Szabò (2016). In the above approches, the credible balls are inflated by a large constant (possibly going to infinity) to ensure that the coverage is correctly bounded from below, so these results although very promising do not completely answer the question of what is the subset $K$ so that (1) is satisfied.

This is an important question, since as said previously statistical analyses only make sense if measures of uncertainty are provided with the inference and since the Bayesian approach naturally provide measures of uncertainty.

## References


Barron, A., Schervish, M., and Wasserman, L. (1999), "The Consistency of Posterior Distributions in Nonparametric Problems," *Ann. Statist.*, 27, 536–561.

Belitser, E., and Ghosal, S. (2016), "Empirical Bayes Oracle Uncertainty Quantification for Linear Regression," Technical report.

Bickel, P. J., and Kleijn, B. J. K. (2012), "The Semiparametric Bernstein–von Mises Theorem," *Ann. Statist.*, 40, 206–237.

Castillo, I. (2010), "A Semiparametric Bernstein-von Mises Theorem for Gaussian Process Priors" *Probability Theory and Related Fields*, 152, 53–99.

Castillo, I., and Nickl, R. (2013), "Nonparametric Bernstein–von Mises Theorems in Gaussian White Noise," *Ann. Statist.*, 41, 1999–2028.

Castillo, I., and Nickl, R. (2014), "On the Bernstein–von Mises Phenomenon for Nonparametric Bayes Procedures," *Ann. Statist.*, 42, 1941–1969.

Castillo, I. and Rousseau, J. (2015), "A General Bernstein–von Mises Theorem in Semi-parametric Models," *Ann. Statist.*, 43, 2353–2383.

Cox, D. D. (1993), "An Analysis of Bayesian Inference for Nonparametric Regression," *Ann. Statist.*, 21, 903–923.

Freedman, D. (1999), "On the Bernstein Von Mises Theorem with Infinite Dimensional Parameter," *Ann. Statist.*, 27, 1119–1140.

Ghosal, S., Ghosh, J. K., and van der Vaart, A. (2000), "Convergence Rates of Posterior Distributions," *Ann. Statist.*, 28, 500–531.

Hoffman, M., Rousseau, J., and Schmidt-Hieber, J. (2015), "On Adaptive Posterior Concentration," *Ann. Statist.*, 43, 2259–2295.

Nurushev, N., and Belitser, E. (2016), "Needles and Straw in a Haystack: Robust Empirical Bayes Confidence for Possibly Sparse Sequences," Technical report.

Rivoirard, V., and Rousseau, J. (2012), "Bernstein–von Mises theorem for Linear Functionals of the Density," *Ann. Statist.*, 40, 1489–1523.

Robert, C. (2001), *The Bayesian Choice*, 2nd ed., New York: Springer.

Rousseau, J., and Szabò, B. T. (2016), "Asymptotic Frequentist Coverage Properties of Bayesian Credible Sets for Sieve Priors in General Settings," Technical report.

Szabó, B. T., van der Vaart, A. W., and van Zanten, J. H. (2015), "Frequentist Coverage of Adaptive Nonparametric Bayesian Credible Sets," *Ann. Statist.*, 43, 1391–1428.

van der Pas, S., SzabB., and van der Vaart, A. (2017), "Uncertainty Quantification for the Horseshoe," Technical report.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge: Cambridge University Press.

# On the Probability of Co-primality of two Natural Numbers Chosen at Random: From Euler identity to Haar Measure on the Ring of Adeles

Sergei Abramovich[†] and Yakov Yu. Nikitin[‡,§]
[†]School of Education and Professional Studies State University of New York, Potsdam, NY, USA
abramovs@potsdam.edu
[‡]Department of Mathematics and Mechanics, Saint Petersburg State University, St. Petersburg, Russia
[§]National Research University—Higher School of Economics, St. Petersburg, Russia
yanikit47@mail.ru

Communicated by Thomas Mikosh

This article overviews the rich history of a truly remarkable problem situated at the confluence of probability theory and theory of numbers—finding the probability of co-primality of two randomly selected natural numbers. The goal is to reveal the genesis of the problem and understand the role of mathematicians of the past in solving this problem and using it as a background for advancing mathematical ideas. The article describes different approaches to the solution of the problem, reviews its various generalizations, examines similar problems, and offers diverse historical perspectives.

## *§1. Introduction*

The problem of computing the probability that two randomly chosen natural numbers are relatively prime (referred to below as the Problem) is one of the gems of mathematical heritage. Its equivalent formulation is finding the probability of the irreducibility of a fraction with randomly selected numerator and denominator from the set of natural numbers. The Problem is quite remarkable despite, or perhaps because of the ease of its formulation, when fractions, which are used as tools in computing chances of simple events already at the primary school level, are themselves considered through the lenses of probability. Yet, the Problem's appeal goes far beyond the elementary level and it may safely be regarded as one of the most famous problems situated at the confluence of the theory of probability and number theory. It can be found in many text/problem-solving books, articles, and monographs. Here are just two examples: in a famous textbook on number theory (Hardy and Wright, 1975) the Problem is presented as Theorem 332; in a more recent book on analytic and probabilistic number theory it is a corollary from Theorem 4 in Tenenbaum (1995, p. 40).

All this motivated the authors to address the following queries. What is the genesis of the Problem? How could it be traced back to the work of others? Who are those others? How can the problem be used to advance our knowledge of the history of mathematics? In this paper, different approaches to the Problem will be described, various generalizations reviewed, reflections on similar problems offered, and some historical perspectives inferred.

## *§2. Two Approaches to Solving the Problem*

### *§2.1. Number Theoretical Approaches*

From the number theory perspective, one can take a large natural number n, then consider all possible pairs of numbers in the range 1 through $n$,

$$B_n = \{(i,k) : 1 \leq i < k \leq n\}$$

and select from them the set of relatively prime pairs, that is, the set

$$A_n = \{(i,k) : 1 \leq i < k \leq n, \mathrm{GCD}(i,k) = 1\}.$$

Consequently, the probability $P$ that two numbers selected at random are relatively prime is quite natural to define as the limit (assuming that it exists)

$$P = \lim_{n\to\infty} P_n := \lim_{n\to\infty} \frac{|A_n|}{|B_n|} = \lim_{n\to\infty} \frac{2|A_n|}{n(n-1)}.$$

That is how the Problem can be solved through a number theory approach. It should be noted though, that the probability is substituted here with the asymptotic density of the set of irreducible pairs that, unlike probability, is not countably additive. Originally, to compute the asymptotic behavior of $|A_n|$, some non-rigorous reasoning was used (an example of which will be given below).

Then, it became clear that $|A_n|$ is equal to the sum of values of Euler phi function $\sum_{k=2}^{n} \varphi(k)$ the asymptotic behavior of which was found first by Dirichlet (1849, 1897) and then (with a better residual term) by Mertens (1874), Kronecker's student, by using the

Möbius function, namely,

$$|A_n| \sim \frac{3}{\pi^2} n^2, \quad n \to \infty.$$

from where, finally, the well-known value $P = 6/\pi^2 \approx 0.607927$ results.

Later, this solution, with gradual simplifications, has been repeated multiple times, e.g., in Sylvester (1883), Kronecker (1901), Hardy and Wright (1975), Apostol (1976), Knuth (1981), and Yaglom and Yaglom (1987).

Another number theoretical approach to the problem is based on the ideas of algebraic number theory. In this approach, the formulation using the *ring of finite integral adeles* $\widehat{\mathbb{Z}}$, which is the well-known compactification of $\mathbb{Z}$, is crucial. $\widehat{\mathbb{Z}}$ is the compact ring densely containing $\mathbb{Z}$ on which there exists a unique Haar probability measure so that all tools provided by probability theory are applicable. The limit theorem solving the Problem becomes then a form of the Law of Large Numbers in the extended probability space, see Kubota and Sugita (2002) and Sugita and Takanobu (2003). This approach was set up by Novoselov (1964) and developed further by many followers (cf. Sugita and Takanobu, 2003) for their list. In the last paper, the reader can find also some promising new advancements.

### *§2.2. A Probabilistic Approach*

Another method of solving the Problem is based on heuristic reasoning when the key step is to use, yet without sufficient justification, the probabilistic property of 'independence' of the events comprised of divisibility of a randomly selected natural number by different prime numbers. After a series of manipulations, the sought probability can be represented through the infinite product $\prod_p (1 - 1/p^2)$ over all prime numbers (see, e.g., Schroeder, 2009, pp. 52–53). Due to the so-called Euler product formula

$$\prod_p (1 - 1/p^2)^{-1} = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \tag{1}$$

discovered by Euler in 1737, once again, the same value $P$ as before results. That was the way many great mathematicians of the past had have approached and solved the Problem, although indicating that their reasoning was not quite rigorous.

Recall that the function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

defined for real and complex values of $s$, is called *zeta function*. For real values of $s$, it was introduced in 1737 by Euler who was able to compute all the values $\zeta(s)$ of the function for even $s$. Notwithstanding, up to this date, the exact values of the function for odd s were unknown, although approximate values could be easily computed to any given accuracy.

In 1859, Riemann studied the zeta function $\zeta(s)$ for complex values of $s$ and revealed the close relationship between the non-trivial zeroes of the function and the distribution of prime numbers. The *Riemann hypothesis*, nowadays being considered as the most famous unsolved problem in mathematics, states that the zeta function does not have non-trivial zeroes outside the line Re$(s) = 1/2$.

Nowadays, the independence of all events related to the divisibility of numbers, previously just discussed, can be carefully proved. Successful examples of using this approach enabling the same answer as a veracious probability can be found, for example, in Tenenbaum (1995) and, especially illustrious, in Pinsky (2014). See also a fabulous book by Kac (1959) devoted to proximal topics.

### *§3. A Russian Perspective*

In Russia, the Problem is often referred to as the "Chebyshev's problem" named after the outstanding Russian mathematician of the 19th century. Chebyshev, as his students reported (e.g. Markov, 1913, p. 173), used to introduce it at his lectures as follows:

> "Find the probability of irreducibility of a rational fraction, the numerator and the denominator of which are chosen at random."

Lyapunov, probably the most famous student of Chebyshev, attended his lectures in 1879–1880 and was known to carefully record them. More than half a century later, Krylov, a notable applied mathematician and the world authority in shipbuilding mechanics, had published Chebyshev's lectures recorded by Lyapunov (Chebyshev, 1936). Excerpts from these lectures translated into English can be found in Sheynin (1994). The lectures include the problem about irreducibility of a randomly selected fraction that Chebyshev solved at the heuristic level of rigor as an example of probabilistic reasoning (Chebyshev, 1936, pp. 152–154). Because, as it follows from Lyapunov's records, Chebyshev, in his lectures, did not mention anybody's name in connection with the Problem, it seems quite natural that his students, Lyapunov included, could have referred to it as "Chebyshev's problem." And because Chebyshev is commonly considered as the grandfather of Russian mathematics, this term appears to be a robust mathematical folklore in Russia.

Known to the authors later references to "Chebyshev's problem" in publications available in Russian include, first of all, the book by Yaglom and Yaglom published originally in 1954 (see problem 90, p. 44).

Note that in its English editions Yaglom and Yaglom (1987) where the Problem is included under the number 92, a reference to Chebyshev is absent. Also, the term "Chebyshev's problem" can be found in several instructional materials on probability and statistics, including Emelyanov and Skitovich (1967), Sveshnikov et al. (1970, problem 4.10, with a reference the 1954 edition of Yagloms), and Zhukov (2004, pp. 157–158).

Alternative evidence, along with reiteration of the Chebyshev's solution, can be found in a textbook by Markov (1913, pp. 173–176), another eminent student of the grandfather of Russian mathematics, widely known for the invention of random processes called Markov chains. Markov notes that the Problem is defined only "after a series of conditions that explain the meaning of the words that the numerator and the denominator of a fraction are chosen at random." Nonetheless, Markov almost neglects the discussion of these conditions and their role in his own reasoning.

It should also be mentioned that heuristic solution by Chebyshev and Markov was seriously criticized by Bernstein (1964, p. 220) who pointed out that in the context of the Problem "one has to determine the limiting or asymptotic behavior of frequencies of natural numbers from a certain class that are distributed according to a certain rule within the sequences under consideration, rather than probabilities which we would never identify experimentally. Those limiting behaviors of ratios represent familiar analogies with mathematical probabilities, and while they are quite important for number theory from a heuristic perspective, mixing these two concepts appears to be an unfortunate mishmash."

### *§4. Evidence Leads to Dirichlet*

In his textbook, in a footnote on p. 173, Markov (1913) mentions that the Problem and its solution can also be found in the lectures by Kronecker (1901), more specifically, in Lecture 24 in which Dirichlet is given precedence for this solution over all other authors. The last remark must be elaborated further. It is known that beginning from 1841, Kronecker studied at the University of Berlin where he attended lectures of Dirichlet and Steiner. In 1845, Kronecker defended a doctoral dissertation on algebraic number theory under the direction of Dirichlet. Soon after the defense, he left Berlin due to family circumstances and returned only in 1855, the year of Gauss's death, when Dirichlet moved to Göttingen to take over Gauss's post. At first, Kronecker did not have an official university position, but, after being elected to Berlin Academy in 1862, started enjoying offers of lectureship appointments.

One can suggest that in his own lectures, Kronecker (1901) reminisced Dirichlet's lectures of the 1841–1843. This brief time span preceded the period when Chebyshev lectured in St Petersburg, something that happened after he succeeded Bunyakovsky at the University of St Petersburg in 1860 (Prudnikov, 1976). Thus, comparing Chebyshev and Dirichlet in the context of the Problem, one can conclude that Dirichlet encountered it earlier then Chebyshev. In addition, in 1852, from June till November, Chebyshev visited Europe and met there several distinguished mathematicians, Dirichlet included Seneta (2001). Who knows, it might be that Chebyshev learned from Dirichlet the problem about the irreducibility of a random fraction.

In Lecture 24, Kronecker attends to the solution of the Problem. He meticulously and with great thorough carries out necessary estimates and proves that the density of pairs of relatively prime numbers over all possible pairs tends to the number $6/\pi^2$. Then, he recalls that Dirichlet in his lectures more often considered the Problem from the probabilistic perspective. This allowed Dirichlet to reach the answer more rapidly, yet his reasoning may not be considered a rigorous proof.

With this in mind, Kronecker analyses Dirichlet's reasoning. Let $w$ be the probability that two randomly chosen natural numbers $i$ and $k$ are relatively prime. Closely connected to $w$ is the probability $w_t$ that the greatest common factor of $i$ and $k$ is equal to $t \geq 1$. It is clear that the original set of pairs is greater than the latter set by the factor $t^2$. Because any pair of natural numbers has a common factor, we have $\sum_{t=1}^{\infty} w_t = 1$ whence

$$w = \frac{1}{\sum_{t=1}^{\infty} \frac{1}{t^2}} = \frac{6}{\pi^2}.$$

Further, Kronecker notes that such an approach to the problem assumes a priori that the probability in question does exist; that is, that there exists the limiting value of the studied ratio of the number of pairs as their total number grows large and that this value is represented in an analytic form. In rigorous terms, Kronecker writes, "what Dirichlet really proves is that the probability in question, if after all it exists, has to be equal to $6/\pi^2$," adding that his own proof is free from this deficiency. By the way, it was Kronecker who served as editor of the posthumous publication of works by Dirichlet (1897).

In his famous paper Dirichlet (1849), that appears also in Dirichlet (1897), Dirichlet solved a complex problem about asymptotic behavior of the sum of the values of the Euler phi function (see Hardy and Wright 2008, p. 359 and Dickson 1952, p. 119). Probably, this allowed Dirichlet to make his reasoning rigorous and to obtain an explicit answer for the Problem. It is quite possible though that in his lectures, Dirichlet simply did not pay attention to details. Nonetheless, Dirichlet did have keen interest in the theory of probability, a fact that a paper by Fischer (1994) supports. In particular, in this paper, a number of aspects of a lecture course on probability theory by Dirichlet have been compared to that of Chebyshev.

### *§5. The Problem's Appeal has Endured Through Time*

In the early 1880s, Cesàro (1881) replicated the Problem (and its solution) from Dirichlet's lectures. In particular, in the first publication, Cesàro very briefly, literally in two lines, formulates a question about the chances of co-primality of just any two natural numbers and states that such chances are 61 to 39. Here, 0.61 is undoubtedly an approximate value of the constant $6/\pi^2$. In the subsequent publications Cesàro (1883, 1884), this statement was justified using arguments similar to those of Dirichlet and, in addition, some further generalizations were provided. Concurrently and independently, a solution to the Problem based on the asymptotic behavior of the sum of the values of Euler phi function had been obtained by Sylvester (1883).

In the modern textbook Bundschuh (2008, p. 52), the author states that solution to the Problem was found by Cesàro and Sylvester and then adds, "it appears that already in the 1849 the solution was found by Dirichlet using somewhat different technique." The above-mentioned testimony of Kronecker, which, Bundschuh, most likely, did not take into account, without a doubt speaks to the favor of Dirichlet as far as the authorship of the Problem is concerned.

The Problem did not escape interest of notable mathematicians of the 20th century, yet, for some reason, without any attention to its rich history. Indicative is a story described by Bellman (1984) in his autobiography. Shortly after the World War II, Bellman, jointly with a number theorist Shapiro, tackled the Problem and the two submitted their solution to the journal *Transactions of American Mathematical Society* which was edited by Kaplansky. The latter sent it to Erdös for review who recommended acceptance. Nonetheless, Kaplansky decided to ask Erdös if he could simplify the proof. Consequently, Erdös found a brief and elegant proof that caused Kaplansky to suggest that the former be a co-author of the submitted paper. But Bellman and Shapiro argued that the length of their proof was not due to the deficiency of the method used but, rather, to its thoroughness and, had they skipped some details, they could have also presented a brief proof. As a result, the paper was withdrawn. From the historical perspective, the essence of this tale is that four eminent mathematicians were not aware of the fact that the problem they discussed has a long history and that a century before them it was posed and solved by Dirichlet, well-known to Chebyshev, and, perhaps, even Gauss had been cognizant of the solution and could have shared it with Dirichlet during one of their get-togethers.

### *§6. Great Gauss Enters the Stage*

One can regard as an improbable speculation the above suggestion that Dirichlet himself might have learned the Problem and its solution from Gauss who, thereby, could have been the first to pose the Problem. Nonetheless, knowing that Dirichlet did meet Gauss several times (in 1827, 1828, and 1849) and enjoyed receiving his letters (Elstrodt, 2007) makes our speculation sound fairly plausible. We will return to the discussion of this hypothesis at the end of the paper.

Furthermore, it is well known that Gauss used not to publish all of his findings. The most famous instances of that include non-Euclidean geometry, quaternions, and the method of least squares. There is a legend (Choi, 2011) that young and talented Jacobi (a close friend of Dirichlet) visited Gauss to show him the newest theory of elliptic functions, something that later has given him the fame. Gauss listened carefully, praised Jacobi, then opened a drawer of his table and pulled out a bunch of notes. Gauss explained to Jacobi that he completed work on this subject some time ago, but did not find the topic worthy of publishing. In turn, puzzled Jacobi asked Gauss as to why he had published even weaker results. An answer to Jacobi's question can be found at the seal of Gauss (Fig. 1) where it is written in Latin "Pauca, sed matura," that is, "Few, but ripe."


Figure 1: Gauss's seal where it is written in Latin "Pauca, sed matura," that is, "Few, but ripe."

### *§7. Equivalent Formulations and Generalization*

There exist related formulations of the Problem worth mentioning. For example, Kranakis and Pocchiola (1994), with a reference to (Knuth, 1981, p. 324) crediting the Problem to Dirichlet, proposed to calculate the probability that in the lattice plane, a segment connecting a randomly chosen lattice point to the origin is free from any other lattice point. It is easy to understand that if the segment does pass through such a point, then the fraction formed by the coordinates of the randomly chosen point is reducible. In a slightly different form, this interpretation is discussed in Apostol (1976, p. 63, Theorem 3.9). Pieprzyk et al. (2013, pp. 190–191) narrowed down the Problem to odd integers and found that the probability of two such numbers, when randomly selected, being co-prime is equal to $8/\pi^2$.

A natural generalization of the Problem is to find the probability/density of $k > 2$ randomly selected natural numbers being relatively prime. The answer, $1/\zeta(k)$, has been known for a long time (see Lehmer (1900) and, for a more modern presentation, Nymann (1972). A similar problem about the probability of a pairwise co-primality of $k$ randomly selected natural numbers has been solved only in the case $k = 3$ with the answer (Finch 2003, p. 110; Schroeder 2009, p. 55):

$$Q = \frac{36}{\pi^4}\prod_p \left(1 - \frac{1}{(p+1)^2}\right) \approx 0.286747.$$

In a remarkable monograph on the distribution of prime numbers, Landau (1909, p. 69) posed the question: What is the probability of randomly selecting a prime number? In accord with the conceptualization of the time, this question was interpreted as computing the corresponding density; that is, the limit $\lim_{x\to\infty} \pi(x)/x$, where $\pi(x)$ is the number of primes not greater than $x$. By that time, Hadamard and Vallee-Poussin had already proved that $\lim_{x\to\infty} \pi(x)\ln(x)/x$ exists and is equal to one (e.g. Hardy and Wright, 1975). Thus, Landau stated that the probability sought is equal to zero. This problem, under the number 94, was included in Yaglom and Yaglom (1987). Chebyshev could have also received the same result using his well-known bilateral approximation of $\pi(x)$; see, for example, Tenenbaum (1995, p. 10).

It is of interest to compute the density/probability of randomly selecting a natural number *free from squares*; that is, not being divisible by any square of a natural number except one. This probability is equal to $1/\zeta(2) = 6/\pi^2$. It was computed by Gegenbauer (1885); see also Hardy and Wright (2008, Theorem 333).

One may wonder, as to why the probabilities of a natural number being free from squares (squarefree) and of two natural numbers being relatively prime coincide. The answer at the physical level of rigor can be found in Schroeder (2009, p. 53). If a natural number $n$ is squarefree, then it may not be divisible by a prime number $p_k$ more than once. In other words, either $n$ is not divisible by $p_k$, or, if is, it would not be divisible for the second time. That is,

$$\Pr(p_k^2 \mid n) = \left(1 - \frac{1}{p_k}\right) + \frac{1}{p_k}\left(1 - \frac{1}{p_k}\right) = \left(1 - \frac{1}{p_k^2}\right).$$

Taking the product of these probabilities over all prime numbers (under the assumption of the independence of the corresponding events) and using identity (1) yields the answer $1/\zeta(2) = 6/\pi^2$.

Even more far-reaching generalization of the squarefree problem is to find the density of numbers free from cubes, fourth powers, and so on, up to the so-called $n$-free numbers; that is, numbers not divisible by any $n$-th power of a natural number. It appears that Gegenbauer (1885) was the first to find that this density is equal to $1/\zeta(n)$. A somewhat more modern presentation of this problem can be found in Evelyn and Linfoot (1931).

In the paper by Hafner et al. (1993), the 'probability' of relative primality of determinants of two randomly chosen $n \times n$ matrices with integer elements is considered. This probability, $\Delta(n)$, is computed and it is expressed in terms of the infinite product,

$$\Delta(n) = \prod_p \left[1 - \left\{1 - \prod_{k=1}^{n}(1 - p^{-k})\right\}^2\right].$$

The value $\lim_{n\to\infty} \Delta(n)$ has been approximately computed by Vardi (1991, p. 174) and is equal to $0.353236$; see also Flajolet and Vardi (1996).

It is possible to extend the Problem to other algebraic structures. Here is an example (cf. Collins and Johnson, 1989, for further details). The ring of Gaussian integers consists of the elements $a + b\mathrm{i}$, where $a$ and $b$ are integers and $\mathrm{i}^2 = -1$. Suppose that two Gaussian integers are chosen at random. The probability/density that they are co-prime, in the sense taken by Collins and Johnson (1989), is equal to $6/(\pi^2 G) \approx 0.663700$, where

$$G = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2} \approx 0.915965,$$

is known as Catalan's constant.

### *§8. Arnold's Perspectives*

It is interesting to recount perspectives on the Problem by Arnold (1937–2010), one of the greatest mathematical minds of the late 20th century, whose opinion bears great weight. Arnold (2003a) is very direct when referring to the problem about irreducibility of a random fraction as a theorem of Gauss. In Arnold (2015, p. 86), it is stated that the "probability was computed by Gauss and the result published by Dirichlet" and Dirichlet (1849) is given as the reference to the above-mentioned result. Moreover, Arnold (2003b, p. 4) hypothesizes that this and similar results could have been known already to Euler, perhaps without proof, something that is not unlikely, and was later completed by Gauss.

In another book, that promotes the notion of mathematics as an experimental science, Arnold (2005) doubled-down on his claim about the authorship of the Problem. To this end, he considers (Fig. 2) 80 lattice points within the disk $\{(x, y) : x^2 + y^2 \leq 25\}$ (excluding the origin), finds that 48 of them have relatively prime coordinates because when such points are connected with the origin, the connecting segment is free from other lattice points (cf. Apostol 1976, p. 63; Kranakis and Pocchiola 1994), and then concludes that "the frequency of irreducibility within the circle is equal to $48/80 = 3/5$, that is, 60%" (Arnold, 2005, p. 13). Arnold argues that through experimenting with larger circles one can compute the limiting probability of reducibility which is approximately 0.608 and then goes on to suggest that "by computing this experimentally found constant, Euler obtained its exact value, $C = 6/\pi^2$. This experimental work led him to a great deal of mathematical discoveries—theory of zeta functions, theory of Fourier series (for rough periodic functions), and to the theory of graduated algebras and their Poincare series" (ibid, p. 14). Furthermore, Arnold believes that "it is due to Euler's experimental investigation of the probability of irreducibility of fractions that identity (1) was developed" (ibid, p. 17). From a historical perspective, all these suggestions give any indication of being incredibly fascinating and terrific, but, unfortunately, the famous author did not provide any references in support of his statements.

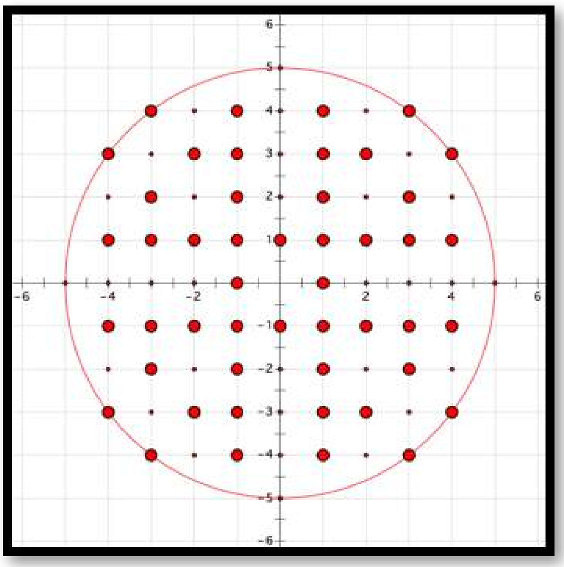



Figure 2: Counting lattice points within the disk $\{(x, y) : x^2 + y^2 \leq 25\}$.

### *§9. Conclusion*

All things considered, the authors have come to conclude as follows. Historical evidence suggests that Cesàro, Chebyshev and Sylvester were not the first to pose and heuristically solve the Problem, although they could have done this independently. Instead, the priority should be given to Dirichlet and, perhaps, to Gauss. Besides, the Problem was revisited multiple times after Dirichlet, Chebyshev, and Cesàro. In particular, as described in Section 5, in mid-20th century, several eminent mathematicians debated various solutions to the Problem without any attention to its rich history. Finally, it would be very interesting to find the traces of the Problem in the vast mathematical heritage of Gauss and Euler, because Arnold (2003b, 2005, 2015) is confident that is the case. This is something that the authors, unfortunately, were not able to do.

**Acknowledgement**. The authors are grateful to Dr. Galina Sinkevich for help and valuable discussions. Research of the second author was supported by grant RFBR No. 16-01-00258 and by grant SPbGU-DFG 6.65.37.2017.

## References

Abrams, A. D., and Paris, M. T. (1992), "The Probability that (a, b) = 1," *Coll. Math. J.*, 23, 47.

Apostol, T. M. (1976), *Introduction to Analytic Number Theory*, New York: Springer.

Arnold, V. I. (2003a), *New Obscurantism and Russian Education*, Moscow: Fazis, In Russian, Available at `http://www.mccme.ru/edu/viarn/obscur.htm`.

Arnold, V. I. (2003b), *Euler Groups and Arithmetic of Arithmetic Progressions*, Moscow: Moscow Center for Continuous Mathematics Education.

Arnold, V. I. (2005), *Experimental Mathematics*, Moscow: Fazis, In Russian.

Arnold, V. I. (2015), *Lectures and Problems: A Gift to Young Mathematicians*, Providence, RI: AMS.

Bellman, R. (1984), *Eye of the Hurricane. An Autobiography*, Singapore: World Scientific.

Bernstein, S. N. (1964/1928), "Contemporary State of the Theory of Probabilities," Collected Works, vol. 4. Moscow: USSR Academy of Sciences Press, In Russian.

Bundschuh, P. (2008), Einfuhrung in die Zahlentheorie, 6 Auflage, Springer-Verlag, In German.

Cesàro, E. (1881), "Question proposee 75," *Mathesis*, 1, 184.

Cesàro, E. (1883), "Question 75 (Solution)," *Mathesis*, 3, 224–225.

Cesàro, E. (1884), "Probabilite de Certains Faits Arithmetiques," *Mathesis*, 4, 150–151.

Chebyshev, P. L. (1936), *Theory of Probabilities*, Lectures delivered in 1879–1880, Published from notes taken by A. M. Lyapunov (A. N. Krylov, ed.), Moscow-Leningrad: USSR Academy of Sciences Press, In Russian.

Choi, Y. (2011, Jan. 25), "Re: Jacobi and Gauss," [Blog comment], Retrieved from `https://ifwisdomwereteachable.wordpress.com/2011/01/25/lesprit-descalier/`.

Collins, G. E., and Johnson, J. R. (1989), "The probability of Relative Primality of Gaussian Integers," In *Symbolic and Algebraic Computation, Proceedings of International Symposium* ISSAC 1988 (P. Gianni, ed.), 252–258. Lecture Notes in Computer Science 358, New York: Springer.

Dickson, L. E. (1952), *History of the Theory of Numbers*, *vol. 1*: *Divisibility and Primality*, New York: Chelsea.

Dirichlet, P. (1849), Uber die Bestimmung der mittleren Werte in der Zahlentheorie, Abhandl. Kgl. Preuss. Acad. Wiss., Berlin, 69–83.

Dirichlet, P. (1897), Werke, Bd. 2. Berlin, 60–64.

Elstrodt, J. (2007), "The Life and Work of Gustav Lejeune Dirichlet (1805–1859)," Analytic Number Theory, Clay Mathematics Proceedings 7, 1–37, Providence, RI: American Mathematical Society.

Emelyanov, G. V., and Skitovich, V. P. (1967), *Problem-Solving Book on Probability and Statistics*, Leningrad: Leningrad University Press, In Russian.

Evelyn, C. J. A., and Linfoot, E. H. (1931), "On a Problem in the Additive Theory of Numbers," *Ann. Math.*, 32, 261–270.

Finch, S. R. (2003), *Mathematical Constants*, New York: Cambridge University Press.
Fischer, H. (1994), "Dirichlet's Contributions to Mathematical Probability Theory," *Historia Mathematica*, 21, 39–63.
Flajolet, P., and Vardi, I. (1996), "Zeta Function Expansions of Classical Constants," Available at http://algo.inria.fr./flajolet/Publications/FlVa96.pdf
Gegenbauer, L. (1885), "Asymptotische Gesetze der Zachlentheorie," Denkschriften Akad. Wiss. Wien, 49, 37–80, In German.
Hafner, J. L., Sarnak, P., and McCurley, K. (1993), "Relatively Prime Values of Polynomials," In Tribute to Emil Grosswald: *Number Theory and Related Analysis* (M. Knopp, M. A. Seingorn, eds.), 437–444, Providence, RI: American Mathematical Society.
Hardy, G. H., and Wright, E. M. (1975), *An Introduction to the Theory of Numbers*, 4th ed., Oxford, UK: Oxford University Press.
Hardy, G. H., and Wright, E. M. (2008), *An Introduction to the Theory of Numbers*, 6th ed., Revised by D. R. Heath-Brown, J. H. Silverman, With a foreword by A. Wiles, Beijing, China: Posts & Telecommunications Press.
Hombas, V. (2013), "What's the Probability of a Rational Ratio being Irreducible?," *Int. J. Math. Ed. Sci. Tech.*, 44, 408–410.
Kac, M. (1959), *Statistical Independence in Probability, Analysis and Number Theory*, The Carus Mathematical Monographs, N 12, Rahway, NJ: The Mathematical Association of America.
Knuth, D. E. (1981), *The Art of Computer Programming: Seminumerical Algorithms*, vol. 2, Reading, MA: Addison-Wesley.
Kranakis, E., and Pocchiola, M. (1994), "Camera Placement in Integer Lattices," *Disc. Comput. Geom.*, 12, 91–104.
Kronecker, L. (1901), *Vorlesungen ueber Mathematik*, Teil 2, Abschnitt 1 (Vorlesungen ueber Zahlentheorie), Bd. 1. Berlin Heidelxsberg GmbH: Springer, In German.
Kubota, H., and Sugita, H. (2002), "Probabilistic Proof of Limit Theorems on Number Theory by Means of Adeles," *Kyushu J. Math.*, 56, 391–404.
Landau, E. (1909), *Handbuch der Lehre von der Verteilung der Primzahlen*, Leipzig und Berlin: BG Teubner, In German.
Lehmer, D. N. (1900), "Asymptotic Evaluation of Certain Totient Sums," *Am. J. Math.*, 22, 293–335.
Markov, A. A. (1913), *Calculus of Probabilities*, 3rd ed., St Petersburg, Russia: Imperial Academy of Sciences Press, In Russian.
Mertens, F. (1874), "Ueber Einige Asymptotische Gesetze der Zahlentheorie," *J. Reine und Angew. Mathem.*, 77, 289–338.
Novoselov, E. V. (1964), "A New Method in Probabilistic Number Theory," *Isv. Ross. Akad. Nauk, Ser. Matem.*, 28, 307–364, In Russian.
Nymann, J. E. (1972), "On the Probability that $k$ Positive Integers are Relatively Prime," *J. Number Theo.*, 4, 469–473.
Pieprzyk, J., Hardjono, T., and Seberry, J. (2013), *Fundamentals of Computer Security*, Berlin: Springer.
Pinsky, R. G. (2014), *Problems from the Discrete to the Continuous. Probability, Number Theory, Graph Theory, and Combinatorics*, New York: Springer.
Prudnikov, V. E. (1976), *Pafnuty Lvovich Chebyshev, 1821–1894*, Moscow: Nauka, In Russian.
Schroeder, M. (2009), *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity*, New York: Springer.
Seneta, E. (2001), "Pafnutii Lvovich Chebyshev (or Tchebichef)," In: C. C. Heyde, E. Seneta, P. Crepel, S. E. Fienberg, J. Gani, editors. *Statisticians of the Centuries*. New York: Springer. pp. 176–180.
Sheynin, O. (1994), "Chebyshev's Lectures on the Theory of Probability," *Arch. Hist. Exact Sci.*, 46, 321–340.
Sugita, H., Takanobu, S. (2003), "The Probability of Two Integers to be Co-prime, Revisited—On the Behavior of CTL-Scaling Limit," *Osaka J. Math.*, 40, 945-976.
Sveshnikov, A. A., Ganin, M. P., Diner, I. Y, Komarov, L. B., Starobin, K. B., Volodin, B. G. (1970), *Problem-Solving Book on Probability, Statistics, and Random Functions*, Moscow: Nauka, In Russian.
Sylvester, J. J. (1883), "Sur le Nombre de Fractions Ordinaires Inegales qu'on peut Exprimer en se Servant de Chiffres qui N'excede Pas un Nombre Donne," C. R. Acad. Sci. Paris, XCVI, pp. 409–413. Reprinted in H. F. Baker (Ed.), The Collected Mathematical Papers of James Joseph Sylvester, vol. 4, Cambridge University Press, 86.
Tenenbaum, G. (1995), *Introduction to Analytic and Probabilistic Number Theory*, New York: Cambridge University Press.
Vardi, I. (1991), *Computational Recreations in Mathematica*, Boston: Addison-Wesley.
Yaglom, A. M., and Yaglom, I. M. (1954/1964/1987), *Challenging Mathematical Problems with Elementary Solutions*, vol. 1: *Combinatorial Analysis and Probability Theory*, New York: Dover, (Russian edition 1954).
Zhukov, A. V. (2004), *Ubiquitous Number*, Moscow: Editorial URSS, In Russian.

# Obituary: Stephen Elliott Fienberg

Stephen Stigler, University of Chicago, US
stigler@uchicago.edu

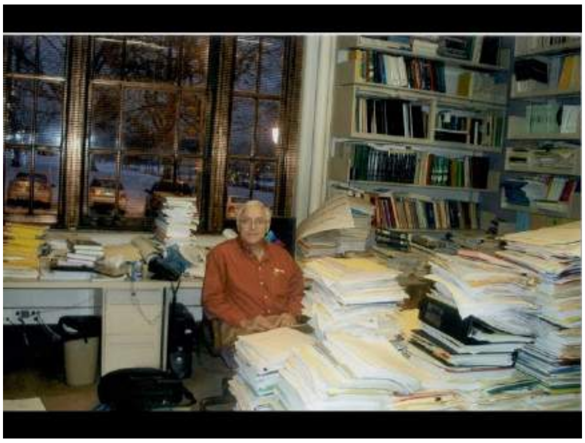
Communicated by the Editor

Stephen Elliott Fienberg: November 27, 1942–December 14, 2016.

Steve Fienberg was born in Toronto, where he remained through his graduation from the University of Toronto in 1964. It was at that University that he first encountered the field of statistics, in a class taught by Don Fraser. The subject proved infectious; he went on to Harvard for his PhD, written under the supervision of Fred Mosteller. At Fred's suggestion, the work of his dissertation was considerably expanded in active partnership with two other Harvard researchers, Yvonne Bishop and Paul Holland, into the very influential book, Discrete Multivariate Analysis, finally published by MIT Press in 1975. That book, colloquially referred to as "Bishop, Fienberg, and Holland" did not invent loglinear models, but it played a crucial role in helping to develop them and inspired a major growth in research in the analysis of categorical data. At a workshop in his honor just two months before he died, Steve told a story that one of his sons was taking a statistics course in college and the instructor approached his son and asked, "Are you any relation to Bishop Fienberg of Holland?" He said the son replied, "I don't believe so, we are Jewish." Steve's first appointment after his PhD was in 1968 at the University of Chicago, where he was jointly in the Department of Statistics and the Department of Theoretical Biology. He left Chicago in 1972 to Chair a new Department of Applied Statistics at the University of Minnesota. In 1980 he moved to Carnegie Mellon's Department of Statistics, where he remained the rest of his life, save for a brief period as Provost at York University in Toronto.

Steve's research developed far beyond his thesis on contingency table models, into network analysis, methodology for confidentiality and of statistical privacy, algebraic statistics, and the application of statistics in science, particularly in social science. He even wrote on the history of statistics, including a paper tracing the history of the term 'Bayesian,' to R. A. Fisher in 1950, who used it in a pejorative sense.

Steve played a major role in what may be called the infrastructure of the profession of statistics. He followed two of his great role models in this, Fred Mosteller and William Kruskal. One part of this was editorial—early on in his career he served as Coordinating and Applications Editor of *JASA*; later he edited the *Annals of Applied Statistics* and was founding or co-founding editor of the *Journal of Privacy and Confidentiality*, *Chance* magazine, and the *Annual Review of Statistics and its Application*. He wrote or co-wrote or edited about 30 books, including *Statistics and the Law* (with Kadane and DeGroot), *Intelligence, Genes, and Success* (with Devlin, Resnick, and Roeder), and *Who Counts?* (on the US Census, with Anderson). But beyond this, Steve played a major role as advisor and critic for many statistical agencies, including the US Census and, over many years, the committees of the National Academy of Sciences, where he played the major role in their 2003 report reviewing the Scientific Evidence on the Polygraph, and was on the National Academy's Report Review Committee, which he co-chaired 2008–2012.



Among Steve's many honors were the 1982 COPPS President's Award, and election to the US National Academy of Sciences, the Royal Society of Canada, and the American Academy of Arts and Sciences. And he did not neglect other aspects of a full life, which for him included ice hockey as coach and player (well beyond the age some of us thought wise, yet with no noticeable loss of teeth). Dinners with Steve and his wife Joyce in great restaurants with fine wines and wide ranging discussion were always a treat. I particularly recall nights in Paris, in Strasbourg, in Dublin, and one night in an Italian restaurant in Manhattan when after we were seated Jackie Kennedy Onassis came in with a friend and sat at the next table, and we realized that the check would likely set a new record (it did).

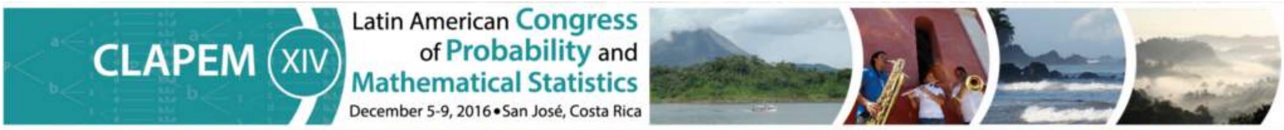
Steve was the senior statesman of statistics in his era, both nationally in the US and internationally. His advisory and editorial activities covered an amazingly broad set of areas and his engagement was always deep and effective. For the past few decades, no meeting or conference on pressing statistical issues such as census undercount or non-reproducibility or ethical experimentation would be complete without his lively and focused participation. Despite his extensive international commitments, he was devoted to his 43 PhD students, and to judge by their comments at the October 2016 meeting at Carnegie Mellon, this devotion was reciprocated. By that time his four-year struggle with cancer was near the end, but no one without that knowledge would have guessed it in view of his vigorous presence throughout the meeting, continuing through the late party (including a small jazz combo) at his home. The energy he brought to his half-century career in statistics seemed undimmed.

**Note from the Editor**: See *Nature* 542, p. 415, for another obituary comment on Steve Fienberg.

# Past Conferences, Meetings and Workshops

## Sponsored and Co-Sponsored by Bernoulli Society for Mathematical Statistics and Probability

## XIV CLAPEM: Dec. 5–9, 2016; San Jose, Costa Rica




The XIV Latin American Congress of Probability and Mathematical Statistics was held at the University of Costa Rica from December 5th to 9th, 2016. The Math School and the Statistics School at the University of Costa Rica were the two main organizers, accompanied by the Math Schools at the Costa Rican Technological Institute and the National University of Costa Rica. The program covered a large variety of thematic sessions, five short courses, contributed sessions and posters contributions. Attendance was around 150 people, and participants came from places as far away as India. From the University of Costa Rica School of Statistics 25 people attended, including professors and, undergraduate and graduate students.

In this occasion we had a very distinguished list of invited speakers, including: Graciela Boente, Universidad de Buenos Aires, Argentina; Alexei Borodin, MIT, USA; Pietro Caputo, Università Roma Tre, Italy; Onesimo Hernandez, CINVESTAV, México; Michael Hudgens from University of North Carolina at Chapel Hill, USA; Susan Murphy from University of Michigan, USA; David Nualart University of Kansas, USA; Gavin Shaddick, University of Bath, UK; and Barry Simon, Caltech, USA.

The School of Statistics at the University of Costa Rica organized a well attended round table with five international experts in Statistics called "ASA statement about p-values: What now?" Panelists were Graciela Boente, Michael Hudgens, Susan Murphy, Gavin Shaddick, and Ricardo Alvarado representing the School of Statistics at the University of Costa Rica. The moderator was Gilbert Brenes, also from that School.

Other highlights from the School of Statistics' participation in CLAPEM included the presentations by professors María Fernanda Alvarado and Mauricio Campos. The titles of their presentations were, respectively, "Considerations on data attributes when modeling Binomial or Poisson grouped data," and "Using Bayesian and traditional meta-analysis methods to study the effects of strength exercises in women with Fibromyalgia."

The Latin American Congress of Probability and Mathematical Statistics (CLAPEM, by its initials in Spanish) is the main event in Probability and Statistics in the region, having been held roughly every two or three years for almost 30 years. It is organized under the auspices of the Bernoulli Society for Mathematical Statistics and Probability and the SLAPEM (Latin-American Society on Probability and Mathematical Statistics). The series of CLAPEMs has greatly contributed to the development of probability and statistics in Latin America by promoting regional cooperation, increasing the scholarly level of the research work in the region, facilitating the collaboration between Latin American researchers and colleagues from the rest of the world.

*Eliana Montero*
*San José*

## PARTY 2017: Jan. 8–13, 2017; Ascona, Switzerland

The aim of the event was to create a friendly atmosphere in which the future generation of PhDs in actuarial science could interact not only with their peers but also with a few well-known experts from actuarial practice and academia. The young researchers had the opportunity to meet researchers not only from their field but also from related disciplines in order to get a flavor of the questions studied and the methods employed.

The actuarial risks discussed at this edition had as main emphasis Ageing and Risk Management. The theme was generous enough to bring forward researchers ranging from life insurance, mortality, longevity and pension funds to risk management, risk measures and risk theory. World-leading researchers in risk management, ageing, population dynamics, pensions, longevity and risk theory (from Switzerland, the UK, France, and Portugal) participated in discussions with practitioners from reinsurance and consulting companies (from the UK, USA and Switzerland) and pension funds (Finland).

The main goal of this scientific exchange was to bridge new synergies across disciplines, and between academia and practice. It succeeded, since this school-conference formula provided the right setting for a first step in future scientific co-operation among very young researchers.

The Winter School managed to bring together 49 young researchers from 19 different countries (Europe, Asia, North America and Africa) in a very friendly atmosphere, helping them to build connections for future collaborations, to network among themselves and with keynote speakers. We like to emphasize that the proportion of women among the participants (young researchers as well as invited speakers) was close to 50%.

The experts from practice, Daria Ossipova from SCOR (Switzerland), Andrew Smith from Deloitte (UK), Jacques Rioux from SAS (USA) and Barbara D'Ambrogi-Ola from Ilmarinen Mutuel Pension Insurance Company (Finland), introduced their latest needs in terms of data analysis and also discussed what it means to be a researcher in industry. The presence of the well-known professors Paul Embrechts (Switzerland), Nicole El Karoui (France), Madhavi Bajekal (UK), Steve Haberman (UK) and Alfredo Egidios dos Reis (Portugal) guaranteed an extensive exchange of information between academia and practice. Bridging academia and practice has been extensively discussed, as the invited speakers reflected upon the importance of research in society.

The Winter School provided the right setting for discussions on a multitude of topics: long term care, financial protection strategies for various groups of insured, investment strategies as supplements for public pensions, life annuities, notional defined contribution pension schemes and other pension products, reinsurance strategies, property and casualty insurance data analysis, uncertainty modeling, population dynamics, socio-economic differences in mortality, old age mortality, stochastic mortality modeling, correlation between financial and demographic factors, distorted risk measures, Parisian ruin, optimal dividend strategies, Solvency II, IBNR and claim reserving, copula models, guaranteed annuity options, Hawkes processes, value at risk, annuitization and machine learning techniques. A few talks presented case studies on real data from Belgium, the Netherlands, Australia, Switzerland and the UK.

This mix-and-match set-up provided the participants with discussions on topics related to their field as well as with a first introduction to the other aspects of (actuarial) risks and ways to address them. More importantly, extensive feedback from key-note speakers or more experienced young researchers were provided to the youngest researchers, who were very enthusiastic about it. All participants have the possibility to publish their research in a special issue of the Risks journal especially devoted to the 2017 PARTY conference. Thanks to the sponsors

The Bernoulli Society has officially endorsed this third edition of PARTY. This successful event would not have taken place without the financial support of the sponsors, to which the organizers and all the participants are very grateful. The generous funding provided by the Congressi Stefano Franscini and ETH Zurich (part of it coming from the Swiss Scientific Research Foundation SNF) was matched by funds from the UK Institute and Faculty of Actuaries, the Swiss Association of Actuaries, Swiss Re, University of Lausanne and Addactis. It allowed the organizers to hold the conference at Monte Verità, a perfect setting for any winter school. It created a unique atmosphere highlighting to the young participants the importance of both scientific as well as cultural exchanges. The latter was amply stressed through an excursion to Bellinzona and a gala dinner in Ascona.

Participants came from all over the world (Norway, Italy, Turkey, Tunisia, Switzerland, UK, France, Kenya, Belgium, Finland, Portugal, Australia, Tunisia, Canada, Sudan, Denmark, China, Japan, USA) due to the affordable registration fee (CHF 575), which included not only the conference fee, but also full board at Monte Verità. Due to the generous contribution of the sponsors, it was also possible to award a few scholarships (registration fee waiver) to participants from Africa. One of the sponsors, Addactis, awarded the best presentation (Julie Thogersen, Aarhus University, Denmark) with financial support to attend the next Insurance: Mathematics and Economics congress in Vienna, July 2017. Additionally, we could reward the next two best presentations (Sarah Kaakai, UPMC, France and Roel Henckaerts, KU Leuven, Belgium) with a refund of their registration fees.

Given the success of this second edition at Monte Verità, a fourth edition of PARTY will probably be held in Romania, in 2019.

We also take pleasure in thanking the staff of Monte Verità for the truly wonderful atmosphere they created throughout the week.

*Séverine Arnold (-Gaille)*
*Lausanne*
*Corina Constantinescu*
*Liverpool*
*Paul Embrechts*
*Zurich*

# Other Events

## Risk Quantification and Extreme Values in Applications: Feb. 15–17, 2017; Lausanne, Switzerland

Following several previous meetings at École Polytechnique Fédérale de Lausanne (EPFL) on risk and extremes, a workshop on *Risk Quantification and Extreme Values in Applications* was held at the Bernoulli Center of the EPFL, from 15–17 February 2017, with financial support from the Chair of Statistics at EPFL and the *Swiss National Science Foundation*. The 12 invited and ten contributed speakers presented their work to an international audience of 85 people containing an unusually high proportion of early-career researchers–indeed the meeting was so successful that it was necessary to find a larger lecture room at the last minute! One day of the scientific programme was devoted to recent advances in statistical methodology for multivariate and spatial extremes, and the other two days focussed on applications and issues that arise in turning this theory into an effective toolkit for risk assessment in areas such as hydrology, meteorology and finance. The event also featured a poster session, with about 25 posters addressing a large variety of topics from the probabilistic properties of extreme value models to new statistical techniques. A speed meeting on the first conference day, which provided a chance for younger researchers to discuss their research with more senior participants, gave a much-appreciated forum for informal discussions and a natural opportunity to network and to foster the exchange of ideas and experiences. Such exchanges were further helped by the main social event, which involved a snowshoe walk to see the sunset over the Lac Léman and copious quantities of speciality local food and drink.

Detailed information on the program and the abstracts of the presentations can be found on the conference web page

http://extremes.epfl.ch/workshop2017

Many of the participants had already attended the two-day winter school on Recent Advances in Extremes prior to the main workshop, which hosted lectures by Clément Dombry and Jonathan Tawn.

*Anthony Davison & Sebastian Engelke*
*Lausanne*

## Modern Problems in Theoretical and Applied Probability: Aug. 22–25, 2016; Akademgorodok, Russia

The Sixth International conference "Modern Problems in Theoretical and Applied Probability" took place from 22 to 25 August 2016 in the Novosibirsk Academy Town (Akademgorodok), the headquarters of the Siberian Branch of the Russian Academy of Sciences. The event was dedicated to the 85-th anniversary of Prof. Alexander A. Borovkov, the founder and leader of the Novosibirsk school of Probability Theory, one of the leading probabilists from the generations of Soviet mathematicians who were directly taught and influenced by A. N. Kolmogorov. The program included about 50 talks by participates from 16 countries around the world, including two 45-minute plenary talks by A. A. Borovkov on "Integro-Local Theorems for Compound Renewal Processes" and I. A .Ibragimov, who talked "On a Problem of Estimation of Infinite-dimensional Parameter in $l_p$ spaces." All the other talks were 30 minute long and given in two parallel sessions over the four days of the conference session talks and were presented participants' results from a spectrum of classical and modern research areas within Probability Theory and its applications. The program also included a poster session accompanied by the short (three minute long) presentations by the young researchers on their posters. The two best presenters from that session were given the opportunity to present 'normal length' talks on their results during the last day of the conference.

The conference was preceded by a satellite Summer School in Advanced Probability that was run from 18 to 21 August jointly by the Novosibirsk State University and Sobolev Institute of Mathematics. The program of the school included two short courses and attracted about 30 under- and postgraduate students interested in probability theory, statistics, and their applications. The lecturers were Prof. P. Morters (University of Bath, UK), who talked about "Branching Processes with Reinforcement," and Dr. M. Lelarge (INRIA, Paris, France), whose course was on "Statistical Mechanics, Graphical Models and Message Passing Algorithms." Following the already existing tradition, the conference was followed by a bus trip to the picturesque Altai Mountains over the following weekend. The list of participants, the titles and abstracts of the talks presented at the conference, and also a large number of photographs can be found at the web page of the conference at

http://math.nsc.ru/LBRT/v1/conf2016

*Kostya Borovkov, Melbourne*
*Sergey Foss, Edinburgh*

# Forthcoming Conferences, Meetings and Workshops, and Calendar of Events

## Sponsored and Co-Sponsored by Bernoulli Society for Mathematical Statistics and Probability


## High Dimensional Statistics, Theory and Practice: October 1–6, 2017; Fréjus, France

This ECAS-SFdS course will be held in October 1–6, 2017 at La Villa Clythia, Fréjus (France) in the French Riviera.

This course, taught in English, presents an introduction and recent advances in high dimensional statistics, with a special emphasis on main concepts of variable selection, nonparametric estimation, supervised and non-supervised classification and multiple testing. It will address theoretical, methodological and practical aspects of this field.

**Website**: `http://ecas2017.sfds.asso.fr`

*Jean-Michel Poggii*
*Orsay*


## Statistics Meets Friends: Nov. 29–Dec. 01, 2017; Göttingen, Germany

On the occasion of Axel Munk's 50th birthday we aim to bridge the gap between Mathematical Statistics, Inverse Problems and Biophysics, highlighting recent developments at their interfaces. Registration for the meeting is open on

`www.stochastik.math.uni-goettingen.de/smf2017`

until July 1st, 2017. The participants will be given the possibility to present a poster.

Keynote speakers include: Rabindra N. Bhattacharya, Lawrence D. Brown, Peter Bühlmann, Tony Cai, Emmanuel Candès, Manfred Denker, Holger Dette, Lutz Dümbgen, Alexander Egner, Klaus Frick, Markus Grasmair, Helmut Grubmüller, Markus Haltmeier, Marc Hoffmann, Chris Holmes, Hajo Holzmann, Thomas Hotz, Zakhar Kabluchko, Bernard A. Mair, Enno Mammen, Richard Nickl, Victor M. Panaretos, Richard Samworth, David O. Siegmund, Alexandre Tsybakov, Sara van de Geer, Aad van der Vaart, and Harrison Huibin Zhou.

*Tatyana Krivobokova*
*Göttingen*

# Other Events

## Symposium on Big data in Finance, Retail and Commerce: Nov. 2–3, 2017; Lisbon, Portugal

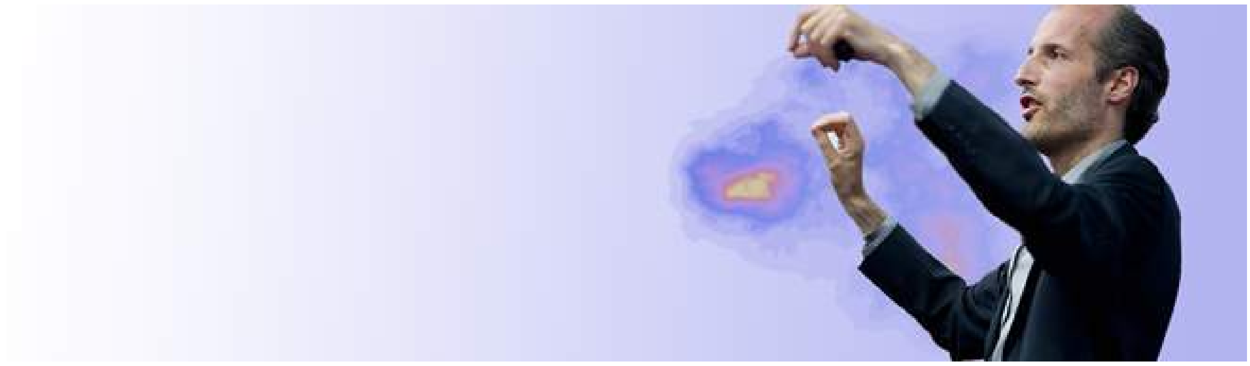
The symposium is organized by the Institute of Financial Big Data of the University Carlos III of Madrid, the Centre of Statistics and its Applications of the University of Lisbon and the Portuguese Statistical Society. The objective of the symposium is to present new approaches to deal with Big Data Applications in finance, retail and commerce bringing together professionals from companies in these areas and researchers who generate sound methods and tools to handle such data. It is hoped that such an interaction may help in understanding the new needs and trends in the emerging field of Data Sciences. Deadline for the submission of abstracts: 16th June 2017.

Further details can be found from

`http://symposiumbigdata2017.weebly.com`

*Feridun Turkman*
*Lisbon*

## Fields Medal Symposium: Oct. 16–19, 2017; Toronto, Canada



The 2017 Fields Medal Symposium will focus on the work of 2014 Fields Medalist Martin Hairer. Martin Hairer's work on regularity structures and rough paths has revolutionized the study of stochastic partial differential equations and related models in statistical physics. The symposium will reflect on this work and will consider its current and potential impact.

The Scientific Program of the Symposium is intended for a broad audience of mathematicians, graduate students, and other scientists. Associated activities include: a Public Opening, featuring a non-technical presentation for a general audience, by Martin Hairer; and a Student Night, involving undergraduates and high school students.

Funding is available for early career researchers, postdocs and graduate students. See the symposium website to apply or register.

Invited speakers include: Gérard Ben Arous, Alexei Borodin, Ajay Chandra, Arnaud Debussche, Hugo Duminil-Copin, Massimiliano Gubinelli, Martin Hairer, David Kelly, Jonathan Mattingly, Jean-Christophe Mourrat, Laure Saint-Raymond, Sylvia Serfaty, Gordon Slade, Herbert Spohn, Hendrik Weber, Lorenzo Zambotti.

The symposium will take place at the Fields Institute; further information can be found from

www.fields.utoronto.ca

*Bryan Eelhart*
*Toronto*

# Calendar of Events

This calendar lists all meetings that have been announced in this and previous issues of *Bernoulli News* together with forthcoming meetings organized under the auspices of the Bernoulli Society or one of its Regional Committees (marked by ).

A more comprehensive calendar of events is available on the ISI Websites

- www.bernoulli-society.org/index.php/meetings
- www.isi-web.org/index.php/activities/calend

**June 2017**

- June 15–17 (2017), *4th International Workshop on Functional and Operatorial Statistics*; Coruña, Spain.
- June 19–23 (2017), *Dynamics, Aging and Universality in Complex Systems*; New York, USA.
- June 25–July 15 (2017), *PCMI Summer Session*; Park City, USA.
- June 26–30 (2017), *10th Extreme Value Analysis (EVA) Conference*; Delft, The Netherlands.
- June 26–30 (2017), *11th Conference on Bayesian Nonparametrics*; Paris, France.

**July 2017**

- July 3–7 (2017), *International Conference on Robust Statistics* (ICORS 2017); Wollongong, Australia.
- July 10–12 (2017), *INFORMS Applied Probability Society Conference 2017*; Evaston, USA.
- July 16–21 (2017), *61st World Statistics Congress*; Marrakesh, Morocco.
- July 17–28 (2017), *Spectral Properties of Large Random Objects*; Bures-sur-Yvette, France.
- July 24–28 (2017), 39th *Conference on Stochastic Processes and their Applications* (SPA); Moscow, Russia.
- July 24–28 (2017), *31st European Meeting of Statisticians*; Helsinki, Finland.

**August 2017**

- August 14–18 (2017), *20th European Young Statisticians Meeting*; Uppsala, Sweden.
- August 25–29 (2017), *XXXIV International Seminar on Stability Problems for Stochastic Models*; Debrecen, Hungary.
- August 28–September 1 (2017), *Dyson–Schwinger Equations, Topological Expansions, and Random Matrices*; New York, USA.

**October 2017**

- October 1–6 (2017), *High Dimensional Statistics, Theory and Practice*; Frejus, France.

**November 2017**

- November 29–December 1 (2017), *Statistics Meets Friends*; Göttingen, Germany.

**February 2018**

- February 27–March 2 (2018), *13th German Probability and Statistics Days*; Freiburg, Germany.

**June 2018**

- June 24–29 (2018), *ISBA 2018 World Meeting*; Edinburgh, UK.

**Quote of the Issue:**

> *"Collaboration is needed, also within the various areas of statistics. Due the complexity of Data Science, cross fertilization is crucial.*[...] *Today, we ARE kindred souls. Times are changing."*

*Sara van de Geer*

# Recent Issues of Official Publications

## Sponsored by Bernoulli Society for Mathematical Statistics and Probability

## *Bernoulli* Vol. 23, No. 3: August 2017

Editor-in-Chief: H. Dette
http://projecteuclid.org/current/euclid.bj

"On Asymptotics of the Discrete Convex LSE of a P.M.F.," F. Balabdaoui, C. Durot & F. Koladjo, 1449–1480.
"Saddlepoint Methods for Conditional Expectations with Applications to Risk Management," S. Kim & K.-K. Kim, 1481–1517.
"Bridge Mixtures of Random Walks on an Abelian Group," G. Conforti & S. Roelly, 1518–1537.
"Predictive Characterization of Mixtures of Markov Chains," S. Fortini & S. Petrone, 1538–1565.
"Tail Asymptotics for the Extremes of Bivariate Gaussian Random Fields," Y. Zhou & Y. Xiao, 1566–1598.
"A Nonparametric Two-sample Hypothesis Testing Problem [...]," M. Tang, A. Athreya, D. L. Sussman, V. Lyzinski & C. E. Priebe, 1599–1630.
"First Time to Exit of a Continuous Itô Process [...]," B. Bouchard, S. Geiss & E. Gobet, 1631–1662.
"Some Theory for Ordinal Embedding," E. Arias-Castro, 1663–1693.
"Constrained Total Undiscounted Continuous-time Markov Decision Processes," X. Guo & Y. Zhang, 1694–1736.
"A General Class of Population-dependent Two-sex Processes with Random Mating," C. Jacob, M. Molina & M. Mota, 1737–1758.
"Universal Scheme for Optimal Search and Stop," S. Nitinawarat & V. V. Veeravalli, 1759–1783.
"Branching Random Walk with Selection at Critical Rate," B. Mallein, 1784–1821.
"Empirical Bayes Posterior Concentration in Sparse High-dimensional Linear Models," R. Martin, R. Mess & S. G. Walker, 1822–1847.
"Probit Transformation for Nonparametric Kernel Estimation of the Copula Density," G. Geenens, A. Charpentier & D. Paindaveine, 1848–1873.
"Efficient Estimation for Diffusions Sampled at High Frequency over a Fixed Time Interval," N. M. Jakobsen & M. Sørensen, 1874–1910.
"Exponential Bounds for the Hypergeometric Distribution," E. Greene & J. A. Wellner, 1911–1950.
"Efficient Particle-based Online Smoothing in General Hidden Markov Models: The PaRIS Algorithm," J. Olsson & J. Westerborn, 1951–1996.
"Quantile Regression for the Single-index Coefficient Model," W. Zhao, H. Lian & H. Liang, 1997–2027.
"Unbiased Simulation of Stochastic Differential Equations using Parametrix Expansions," P. Andersson & A. Kohatsu-Higa, 2028–2057.
"Representations for the Decay Parameter of Markov Chains," J. Chen, S. Jian & H. Li, 2058–2082.
"Behavior of the Wasserstein Distance Between the Empirical and the Marginal Distributions of [...]," J. Dedecker & F. Merlevède, 2083–2127.

## *Stochastic Processes and their Applications* Vol. 127, No. 5: May 2017

Editor-in-Chief: H. Dehling
http://www.sciencedirect.com/science/journal/03044149

"Two–Parameter Process Limits for an Infinite–server Queue with Arrival Dependent Service Times," G. Pang & Y. Zhou, 1375–1416.
"Multilevel Sequential Monte Carlo Samplers," A. Beskos, A. Jasra, K. Law, R. Tempone & Yan Zhou, 1417–1440.
"Constrained BSDEs Representation of the Value Function in Optimal Control [...]" E. Bandini & M. Fuhrman, 1441–1474.
"Least Squares Estimators for Stochastic Differential Equations Driven by Small Levy Noises", H. Long, C. Ma & Y. Shimizu, 1475–1495.
"Finite Dimensional Fokker–Planck Equations for Continuous Time Random Walk Limits," O. Busani, 1496–1516.
"Weak Convergence of the Empirical Truncated Distribution Function of the Levy Measure of [...]," M. Hoffmann & M. Vetter, 1517–1543.
"A Random Cell Splitting Scheme on the Sphere," C. Deuss, J. Hörrmann & C. Thäle, 1544–1564.
"Change of Measure up to a Random Time: Details," D. Kreher, 1565–1598.
"Multidimensional Lévy White Noise in Weighted Besov Spaces," J. Fageot, A. Fallah & M. Unser, 1599–1621.
"Normal Approximation and Almost sure Central Limit Theorem for Non–Symmetric Rademacher Functionals," G. Zheng, 1622–1636.
"New Deviation Inequalities for Martingales with Bounded Increments," E. Rio, 1637–1648.
"Tail Generating Functions for Extendable Branching Processes," S. Sagitov, 1649–1675.
"On the Limiting Law of the Length of the Longest Common and Increasing [...]," J.-C. Breton, C. Houdre, 1676–1720.

## *Bernoulli Society Bulletin e-Briefs* Vol. 29: April 2017

Editor-in-Chief: C. Constantinescu
http://goo.gl/G9AOgl

## Co-Sponsored by Bernoulli Society for Mathematical Statistics and Probability

Have a look at http://goo.gl/7EP2cZ for the latest articles in *Electronic Communications in Probability*, *Electronic Journal of Probability*, *Electronic Journal of Statistics*, *Probability Surveys and Statistics Surveys*, as well as *International Statistical Review*.

# Who is Who in the Bernoulli Society

| | | |
|---|---|---|
| **Executive Committee 2015–2017** | | |
| *President* | Sara van de Geer (Switzerland) | geer@stat.math.ethz.ch |
| *Past President* | Wilfrid Kendall (UK) | w.s.kendall@warwick.ac.uk |
| *President Elect* | Susan Murphy (USA) | samurphy@umich.edu |
| *ISI Director* | Ada van Krimpen (Netherlands) | an.vankrimpen@cbs.nl |
| *Membership Secretary* | Mark Podolskij (Denmark) | mpodolskij@math.au.dk |
| *Publications Secretary* | Thomas Mikosch (Denmark) | mikosch@math.ku.dk |
| *Scientific Secretary* | Byeong U. Park (South Korea) | bupark2000@gmail.com |
| *Treasurer* | Lynne Billard (USA) | lynneb@uga.edu |
| **Council Member 2013–2017** | | |
| | Alicia Carriquiry (USA) | alicia@iastate.edu |
| | Jeng-Min Chiou (Taiwan) (2015–2017) | jmchiou@stat.sinica.edu.tw |
| | Jana Jureckova (Czech Republic) | jurecko@karlin.mff.cuni.cz |
| | Pedro Mendez (Costa Rica) | Pedro.Mendez@ucr.ac.cr |
| | Byeong U. Park (South Korea) (2013–2015) | bupark2000@gmail.com |
| | B. L. S. Prakasa Rao (India) | blsprao@gmail.com |
| | Michael Sorensen (Denmark) | michael@math.ku.dk |
| **Council Member 2015–2019** | | |
| | Arup Bose (India) | bosearu@gmail.com |
| | Valerie Isham (UK) | v.isham@ucl.ac.uk |
| | Carenne Ludeña (Venezuela) | carinludena@gmail.com |
| | Victor Rivero (Mexico) | rivero@cimat.mx |
| | Akira Sakai (Japan) | sakai@math.sci.hokudai.ac.jp |
| | Richard Samworth (UK) | r.samworth@statslab.cam.ac.uk |
| | Lorenzo Zambotti (France) | lorenzo.zambotti@upmc.fr |
| | Johanna Ziegel (Switzerland) | johanna.ziegel@stat.unibe.ch |
| **Committee Chairs** | | |
| *Conferences on Stochastic Processes* | Kavita Ramanan (USA) | kavita_ramanan@brown.edu |
| *Probability and Statistics in the Physical Sciences* | Konstantin Zuev (USA) | kostia@caltech.edu |
| *Publications Committee* | Thomas Mikosch (Denmark) | mikosch@math.ku.dk |
| **Regional Committee Chairs** | | |
| *European* | Niels Richard Hansen (Denmark) | niels.r.hansen@math.ku.dk |
| *East-Asian and Pacific* | (in process of being reformed) | |
| *Latin America* | Carenne Ludeña (Venezuela) | carinludena@gmail.com |
| **Editors** | | |
| *Bernoulli* | Holger Dette (Germany) | holger.dette@ruhr-uni-bochum.de |
| *Stochastic Processes and their Applications* | Herold Dehling (Germany) | herold.dehling@ruhr-uni-bochum.de |
| *International Statistical Review* | Nalini Ravishanker (USA) | nalini.ravishanker@uconn.edu |
| | Ray Chambers (Australia) | ray@uow.edu.au |
| *Bernoulli News* | Miguel de Carvalho (UK) | miguel.decarvalho@ed.ac.uk |
| *Bernoulli e-Briefs* | Corina Constantinescu (Liverpool, UK) | corina.constantinescu@liverpool.ac.uk |
| **Web Editors** | | |
| *Bernoulli Society* | Erik Baurdoux (UK) | e.j.baurdoux@lse.ac.uk |
| *Bernoulli Journal / Bernoulli News* | Justin van der Veeke (Netherlands) | isiwebmaster@yahoo.com |
| *SLAPEM* | Claudio Landim (Brazil) | landim@impa.br |
| *Site Administrator* | Björn Böttcher (Germany) | bjoern.boettcher@tu-dresden.de |
| **Representatives** | | |
| *World of Statistics Committee* | Amber Puha (USA) | apuha@csusm.edu |
| *Bernoulli Youth* | Corina Constantinescu (UK) | C.Constantinescu@liverpool.ac.uk |