

## POISSON APPROXIMATION

By A. D. BARBOUR, L. HOLST and S. JANSON: 280 pp., £30.00, ISBN 0 19 852235 5 (Oxford University Press, 1992).

Discrete probability is coming back into fashion, under 'applied' influences ranging from computer algorithms to molecular genetics, after a couple of generations of being regarded as a backwater by mathematical probabilists. This monograph is the first to treat one of the most interesting newly-developed techniques, the Stein-Chen method for establishing explicit upper bounds for the Poisson approximation. Consider events  $A_1, \dots, A_n$  with  $P(A_i) = p_i$ , and let  $X$  be the random number of events which occur, so  $X$  has expectation  $\lambda = \sum_i p_i$ . It has long been realized that if  $n$  is large, the individual  $p_i$  are small and the events are 'pretty much independent', then the distribution of  $X$  should be approximately the Poisson( $\lambda$ ) distribution. Classically, this idea is formalized by limit theorems proved using transform or method-of-moments techniques. The Stein-Chen method improves classical results by giving bounds for finite  $n$ , and indeed these often lead to better limit theorems. The starting point for this book is a certain coupling lemma, going back to [5]. To say the idea in words, consider the example where  $X$  is the number of empty boxes when  $b$  balls are thrown at random into  $n$  boxes. So  $A_i$  is the event 'box  $i$  is empty'. Now fix one box, say box  $j$ , and let  $X_j$  be the number of empty boxes (excluding  $j$ ) when we condition on box  $j$  being empty. The coupling lemma says we can upper bound the error in the Poisson approximation for  $X$  in terms of the sum  $\sum_j E|X - X_j|$ . And it is easy to bound the summands, because the distribution of balls in boxes conditional on box  $j$  being empty can be obtained from the unconditional distribution by simply re-throwing the balls which happened to land in box  $j$ .

Once you see how this example works, it is clear that a wide range of classical combinatorial probability problems involving urn models, random permutations and random graphs should be do-able in the same way. Much of the book is devoted to working carefully through these examples, and related classical topics such as spacings and maxima of stationary processes, and developing variations of the basic coupling lemma which give sharp bounds in different settings. The final chapter deals with the more recent ideas of Poisson *process* approximations.

The book is well written, and thorough in its coverage of the chosen topics. It is a rare example of a book that is both examples-oriented and rigorous, and it deserves to be read by everyone interested in discrete probability. Perhaps the only quibble I have is that an alternative implementation of the method (which the authors call the 'local approach', developed in, for example, [3, 4]) is mentioned but not treated in detail, so it is hard for the reader to judge the merits of the two implementations.

Let me end with two comments on the big picture.

1. From the theoretical viewpoint, the major open question is to what extent the Stein-Chen technique is useful in the broader world of Poisson and compound Poisson approximation discussed heuristically in [1]. A specific 'challenge problem' concerns the compound Poisson approximation for sojourn times of a continuous-time Markov chain in a rare region; [2] gives the rather unsatisfactory results obtainable with current techniques.

2. To a theoretician, techniques for obtaining sharper results are self-evidently worthwhile. But there is a lot of misguided work done under the 'applied' label,

D. G. VASSILIEV

consisting of detailed studies of models which are very crude simplifications of reality (for example, average case analysis of graph algorithms assuming the graph 'random graph' model, or DNA sequence-matching assuming a Markov chain model). For most applied work, crude results which are very robust under changes in assumptions are more meaningful. In this sense, powerful techniques like the Stein-Chen method can be dangerous in applications, because they focus on strengthening conclusions rather than on making models more realistic.

References

1. D. J. ALDOUS, *Probability approximations via the Poisson clumping heuristic* (Springer, Berlin, 1989).
2. D. J. ALDOUS and M. BROWN, 'Inequalities for rare events in time-reversible Markov chains II', *Stochastic Process. Appl.*, to appear.
3. R. ARRATIA, L. GOLDSTEIN and L. GORDON, 'Two moments suffice for Poisson approximations: the Chen-Stein method', *Ann. Probab.* 17 (1989) 9-25.
4. R. ARRATIA, L. GOLDSTEIN and L. GORDON, 'Poisson approximation and the Chen-Stein method', *Statist. Sci.* 5 (1990) 403-434.
5. A. D. BARBOUR and G. K. EAGLESON, 'Poisson approximation for some statistics based on exchangeable trials', *Adv. in Appl. Probab.* 15 (1983) 585-600.

D. J. ALDOUS

DIRICHLET FORMS AND ANALYSIS ON WIENER SPACE

By NICOLAS BOULEAU and FRANCIS HIRSCH: 326 pp., US\$69.00, ISBN 3 11 012919 1 (Walter de Gruyter, 1991).

We have here an immaculately organised account of Malliavin calculus in the language of Dirichlet forms. The authors show that this is a good language by building up the structure required for Malliavin calculus step by step, and examining carefully what can be said at each stage. The emphasis is on making the weakest hypothesis that is necessary for each result.

For a symmetric Markov process  $X_t$  with invariant measure  $m$ , the associated Dirichlet form is defined by

$$\mathcal{E}(f) = \lim_{t \downarrow 0} \left( \frac{f - P_t f}{t}, f \right)_{L^2(m)} = - \frac{d}{dt} \Big|_{t=0} \mathbb{E}_m f(X_t) f(X_0)$$

whenever the limit exists. When one studies the behaviour of  $X_t$  in equilibrium, this provides a tool which is more subtle than the infinitesimal generator. In particular, the domain of  $\mathcal{E}$  is closed under composition with Lipschitz functions.

Malliavin calculus is based on the Ornstein-Uhlenbeck process on Wiener space. This is a Markov process on  $C(\mathbb{R}^+, \mathbb{R})$ , symmetric with respect to Wiener measure  $m$ , whose Dirichlet form is given by

$$\mathcal{E}(f) = \frac{1}{2} \sum_{n=0}^{\infty} n \|f_n\|_{L^2(m)}^2,$$

where  $f = \sum_{n=0}^{\infty} f_n$  in the Wiener chaos expansion.

For an orthonormal set  $h_1, \dots, h_k$  in  $L^2(\mathbb{R}^+)$ , and  $\phi$  in  $C^2(\mathbb{R}^k)$  of polynomial growth,

consider  $f = \phi(h_1, \dots, h_k)$ , where  $\mathcal{E}(f)$  then we have

$$\mathcal{E}(f)$$

This characterises  $\mathcal{E}$  and shows in  $\mathbb{R}^k$ .

Both Dirichlet forms are proliferating. Although the process is easy, to develop the general theory a lot of machinery. This book learning the highly-developed organisation to the reader's in sooner or later!

The question is then whether rewards justify the efforts? A powerful structure on which processes in equilibrium. While technical difficulties, but the standard reference is Fukushima (North-Holland, 1980). The contribution to the literature, state space, and collecting the operator, absolute continuity also the forthcoming book of

Whereas appreciation of Malliavin's papers of the 1970s explosive. Excitement was gained new depth in probability theory probabilistic treatment of Hilbert transition density of a diffusion result is not covered by Bouleau absolute continuity of the law do not treat the question of further input from Dirichlet Malliavin calculus carefully: versions of the same ideas.

In contrast to general theory Malliavin calculus until much of a stochastic differential rewards, which are treated in and the theory of capacity or to quasi everywhere' refine Malliavin calculus has provided for those of us who enjoy a on the ground of application