Asymptotic Expansions for Statistics Computed from Spatial Data

P. GARCÍA-SOIDÁN

Departamento de Estadística e Investigación Operativa,
Universidad de Vigo, Torrecedeira 105, 36208, Vigo, Spain

SUMMARY

The Edgeworth expansions for dependent data are generalized to the context of spatial patterns, with the aim of obtaining asymptotic expansions which approximate the distribution of statistics computed from spatial data, generated by a weakly dependent coverage process. In particular, the case of estimating the expected proportion (its porosity) of a region that is not covered by the process is treated in detail and explicit formulae are given in the context of a Boolean model, assuming that the random sets generating the model are essentially bounded and satisfy a version of Cramér's condition.

Keywords: BOOLEAN MODEL; COVERAGE PROCESS; CRAMÉR'S CONDITION; EDGECWORTH EXPANSION; POROSITY; VACANCY.

1. INTRODUCTION

The aim of this paper is to investigate the validity of the Edgeworth expansion techniques for spatial patterns, with the purpose of obtaining asymptotic expansions that approximate the distribution of statistics defined in this context.

We assume that the coverage process is weakly dependent in the sense that it has exponentially decaying high-order correlations. Theoretical limit approximations may then be obtained via the central limit theorem or the law of large numbers; see, for example, Hall (1988) in the case of estimating the expected vacancy.
In this paper, the work of Götze and Hipp (1983) on dependent
data is generalized to the context of spatial patterns. Thus, in Section
2, we prove the validity of higher order approximations in developing
distribution of statistics computed from spatial data generated by a
weakly dependent stochastic process.

We aim to adapt the idea of working with a conditional Cramér
condition, given in Götze and Hipp (1983), in the sense that the problem
of checking an adequate conditional Cramér condition may be reduced
to that of proving the standard smoothness condition imposed.

Thus, both non-Studentized and Studentized statistics are consid-
ered and the arguments of Bhattacharya and Rat (1976) (for sums of
i.i.d. random vectors) and those of Bhattacharya and Ghosh (1978) (for
functions of sums of i.i.d. random vectors) are generalized to the con-
text of dependent data. The formal construction of these expansions
may be obtained either by expanding the characteristic function, for the
non-Studentized statistic, or from a combinatorial formula giving the
order of the cumulants, for the Studentized estimator. This is analogous
to the approach in Hall (1992) when the random vectors involved are
independent.

We are particularly concerned with the problem of estimating the
expected proportion of a region that is not covered by the process (see
Hall, 1988) and, for the purposes of this paper, we consider the case
when the stochastic process is generated by a Boolean model. Thus, in
Section 3, we give explicit formulae for approximating the distribution
of these statistics, assuming that the random sets generating the model
are essentially bounded (although we may relax this hypothesis) and
satisfy a version of Cramér's condition.

2. NOTATION AND MAIN RESULTS

In this section, we introduce some important definitions and notation,
beginning with the concept of a coverage process.

Let \( \{\xi_i : i \geq 1\} \) be a stochastic process and \( \{S^{(i)} : i \geq 1\} \) be a
sequence of i.i.d. random sets independent of the process. Then, the
collection of random sets \( P = \{\xi_i + S^{(i)} : i \geq 1\} \) is called a coverage
process. In particular, when the stochastic process is a stationary Poisson
process in \( \mathbb{R}^k \) and \( S^{(i)}, i \geq 1 \), are i.i.d. random subsets of \( \mathbb{R}^k \) this
is referred to as a Boolean model in the continuum \( \mathbb{R}^k \); thus, we say