EDGEWORTH EXPANSIONS AND BOOTSTRAP FOR DEGENERATE VON MISES STATISTICS

F. Götze[1] R.Zitikis[2]

Fakultät für Mathematik
Universität Bielefeld

Institute of Mathematics and
Informatics, Vilnius

December 1993

Abstract. We prove Edgeworth expansions for degenerate von Mises statistics like the Beran, Watson, and Cramér-von Mises goodness-of-fit statistics. Furthermore, we show that the bootstrap approximation works up to an error of order $O(N^{-1/2})$ and that Bootstrap based confidence regions attain a prescribed confidence level up to the order $O(n^{-1})$.

1. Introduction, main results, and some examples

Let $\{\Omega,\, \mathcal{A},\, \mathbb{P}\}$ be a probability space, and let $\mathbf{X} : \{\Omega,\, \mathcal{A}\} \to \{\Lambda, \mathcal{F}\}$ be a random variable. The distribution function of $\mathbf{X}$ will be denoted by $\mathbb{F}$. Furthermore, we shall denote by $\mathbf{X}_1, \mathbf{X}_2, \ldots$ independent copies of $\mathbf{X}$.

In this paper we consider various approximations for the (degenerate) von Mises statistic

$$V_N \triangleq \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{N} H(\mathbf{X}_j, \mathbf{X}_k),$$

[1]Research supported by the SFB 343 at Bielefeld.
[2]Research supported by the SFB 343 and Graduiertenkolleg at the University of Bielefeld.
1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.
Key words and phrases. Bootstrap, coverage probabilities, confidence intervals, Edgeworth expansions, von Mises statistics, U-statistics.

where $H : \Lambda \times \Lambda \to \mathbb{R}$ is a symmetric measurable function which is assumed to be degenerate with respect to $\mathbb{P}$, that is $\mathbb{E}\big(H(\mathbf{X}_1, \mathbf{X}_2)|\mathbf{X}_2\big) = 0\,\mathbb{P}$-a.s. Furthermore, we assume $\mathbb{E}\big|H(\mathbf{X}, \mathbf{X})\big| + \mathbb{E}H^2(\mathbf{X}_1, \mathbf{X}_2) < +\infty$. We use the notation $P_N(x) \triangleq \mathbb{P}\{V_N \leq x\}$.

Our goal is to investigate the following three closely related problems for the statistic V_N:

1) A bootstrap approximation;

2) The bootstrap based coverage probabilities;

3) Edgeworth expansions.

The bootstrap version of V_N we use in this paper is defined (see Bretagnolle (1983)) as follows:

$$V_n^* \triangleq \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} H^*(\mathbf{X}_j^*, \mathbf{X}_k^*),$$

$$H^*(\mathbf{x}, \mathbf{y}) \triangleq H(\mathbf{x}, \mathbf{y}) - \mathbb{E}^* H(\mathbf{x}, \mathbf{X}_2^*) - \mathbb{E}^* H(\mathbf{X}_1^*, \mathbf{y}) - \mathbb{E}^* H(\mathbf{X}_1^*, \mathbf{X}_2^*),$$

where $\mathbb{E}^*$ denotes the expectation with respect to the empirical distribution function $\mathbb{F}^*$ given the sample $\mathbf{X}_1, ..., \mathbf{X}_N$.

We shall use the notation $P_n^*(x) \triangleq \mathbb{P}^*\{V_n^* \leq x\}$.

The limit distribution of V_N, say P_∞, was described by von Mises (1947), and it is the distribution of the random variable

$$V_\infty \triangleq \mathbb{E}H(\mathbf{X}, \mathbf{X}) + \sum_{k=1}^{\infty} \lambda_k (G_k^2 - 1),$$

where $\lambda_1, \lambda_2, \ldots, |\lambda_1| \geq |\lambda_2| \geq \ldots$ are eigenvalues of the self–adjoint Hilbert–Schmidt operator $S : L_2(\Lambda, \mathbb{F}) \to L_2(\Lambda, \mathbb{F})$ defined by

$$Sf \triangleq \mathbb{E}H(\cdot, \mathbf{X})f(\mathbf{X}) = \int H(\cdot, \mathbf{y})f(\mathbf{y})\mathbb{F}(d\mathbf{y}).$$

Here $G_1, G_2, \ldots$ are independent copies of a Gaussian random variable G with mean 0 and variance 1. See Gregory (1977), Rubin and Vitale (1980), Serfling (1980), Dynkin and Mandelbaum (1983) for more details on this representation.

In the same way we may describe the distribution function P_∞^* of the random variable V_∞^* ($\triangleq$ the weak limit, as $n \to +\infty$, of the statistic V_n^* when the random variables $\mathbf{X}_1, ..., \mathbf{X}_N$ are fixed).

If it is not stated otherwise, we shall always assume that the following holds:

The operator S has an infinite number of non-zero eigenvalues λ_j and the random variables $H(\mathbf{X}, \mathbf{X})$ and $H(\mathbf{X}_1, \mathbf{X}_2)$ have moments of all orders.