AN APPROXIMATE CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN QUANTILES IN A BIO-MEDICAL PROBLEM

by

Willem Albers * and Peter Löhnberg †

Abstract

In this note it is described how two distributions arising in a bio-medical investigation are compared by means of a confidence interval for the difference of appropriate quantiles. It is briefly indicated how an approximation for such a confidence interval is derived.

Key words & Phrases: confidence, interval, quantiles, distribution-free, asymptotics

1. Introduction

A surgical lesion in a certain brain structure is known to decrease pathological tremor. Before surgery, neuronal activity has been recorded by means of a microelectrode at several sites in this structure for several patients. Hand acceleration was recorded simultaneously. Tremor was characterized by two measures: momentary acceleration as well as its envelope. The relation between neuronal activity, tremor and microelectrode recording was represented by the same stochastic model structure for both measures of tremor. This model structure was verified from all recordings together (Löhnberg & Hoogstraten). The model indicated how the relative contribution $R$ by the neuronal activity near the electrode tip to the tremor power can be approximated from a recording (Löhnberg, b). This yielded an experimental estimate population $<\hat{R}_E>$ for each tremor measure.

For improved surgery and for insight into the tremor generation mechanism, it was relevant to derive from this population for each tremor measure whether $R$ could be assumed to be 0 for all sites. If nonzero values of $R$ would occur, it was expected that these values would be substantially larger than 0. Some information about these possible nonzero values of $R$ was appreciated. Statistical properties of $\hat{R}$ have been approximated analytically under a number of assumptions (Löhnberg, a). As in this case not all of these assumptions were valid, it was decided to simulate both models for various choices of parameter $R$ including zero. Values of other model parameters were derived from all measurements together. The simulations showed as expected that for $R > 0$, the estimates $\hat{R}$ tended to be stochastically larger than for $R = 0$. Hence it appeared relevant to investigate, to what extent the

* Dept. of Appl. Mathematics, Twente University of Technology
† Dept. of Electrical Eng., Twente University of Technology
distribution of each $<\hat{R}_E>$ contained higher right tail probabilities than the distribution of the estimates $<\hat{R}_0>$ from the related simulations for $R = 0$.

To obtain some insight in this matter, a comparison was made between the highest quantiles of $<\hat{R}_E>$ and $<\hat{R}_0>$ which left the minimally required 5 expected observations in the upper cell of the contingency table (Conover, 1971). Equality was accepted for momentary acceleration and rejected for the envelope. In order to indicate the relevance (Molenaar, 1977) of the difference between the quantiles, its confidence interval was estimated. Since no results seemed to be readily available on (asymptotic) confidence intervals for the difference between two quantiles, we supply some detail on the derivation in the next section. These intervals confirmed the tests, using more properties of the sample distributions.

2. The approximate confidence interval

Let $X_1, \ldots, X_m$ be a sample from a distribution function (df) $F$. For $0 < p < 1$, let $\xi_p = F^{-1}(p)$ be the $p$'th quantile of $F$. We shall assume that $F$ has a density $f$ in a neighbourhood of $\xi_p$ which is positive and continuous at $\xi_p$. Then it is well-known, see e.g., Serfling (1980), p.77, that

$$\hat{\xi}_{pm} = AN(\xi_p, \frac{p(1-p)}{mf^2(\xi_p)})$$

(2.1)

where $\hat{\xi}_{pm} = F_m^{-1}(p)$ is the sample $p$'th quantile and $AN$ stands for asymptotically normal.

Now suppose that independent of the first sample we have a second sample $Y_1, \ldots, Y_n$ from a df $G$ satisfying similar smoothness conditions as $F$. Then in analogy to (2.1) we know that its sample $p$'th quantile, say $\hat{\eta}_{pn}$, is $AN(\eta_p, (p(1-p)) / (ng^2(\eta_p)))$, where $\eta_p = G^{-1}(p)$.

From the above it readily follows that an approximate confidence interval with confidence coefficient $1 - 2\alpha$ for

$$d = \xi_p - \eta_p$$

(2.2)

is given by

$$|d - (\hat{\xi}_{pm} - \hat{\eta}_{pn})| \leq K_\alpha \left[ \frac{p(1-p)}{mf^2(\xi_p)} + \frac{p(1-p)}{ng^2(\eta_p)} \right]^{1/2},$$

(2.3)

where $K_\alpha$ is the $(1 - \alpha)'$th quantile of the standard normal distribution (cf. Serfling (1980), p.103, who gives the result for a single quantile).

A remaining complication is that application of (2.3) is only possible if $f$ and $g$ are known. If this is not the case a modification can be used which replaces these quantities by consistent estimates (see again Serfling (1980) for the one-sample case). In fact, let $X_{m1} < \cdots < X_{mm}$ denote the order statistics of $X_1, \ldots, X_m$ and let $\{k^-_m\}$ and $\{k^+_m\}$ be sequences of integers such that
\[ \frac{k_m^\pm}{m} = p \pm K_\alpha \left\{ \frac{p(1-p)}{m} \right\}^{1/2} + o(m^{-1/2}) \]

Then it follows that (cf. Serfling (1980), p.104)

\[ X_{mk} - \hat{\xi}_{pm} = \pm K_\alpha \left\{ \frac{p(1-p)}{mf^2(\xi_p)} \right\}^{1/2} + o(m^{-1/2}) \quad a.s. \quad (2.4) \]

Applying this result and a similar one for the second sample to (2.3) we obtain the desired approximate interval

\[ |d - (\hat{\xi}_{pm} - \hat{\eta}_{pn})| \leq \frac{1}{4} \left\{ (X_{mk} - X_{mk})^2 + (Y_{nk} - Y_{nk})^2 \right\}^{1/2}. \]

References


Löhnberg, P. (b), Minimal contribution by sources of one variable to the power of another, revised for IEEE Transactions on Biomedical Engineering.

Löhnberg, P. & L.L.W. Hoogstraten, Estimated relative contribution by ventrolateral thalamic activity to pathological tremor power, Report 304, Bio-information group, THT.


Received October 1982, revised June 1983.