BULL. AUSTRAL. MATH. SOC. 62H25
VOL. 12 (1975), 155-157.

# Abstract of Australasian PhD thesis

## Generalised partial correlation and principal components

**Neville Robert Bartlett**

Basically, the thesis consists of three main parts that can be briefly summarised as

(i) introduction and preliminaries,

(ii) general partial correlation and applications, and

(iii) optimality of principal components.

Chapters 1 and 2 contain the introduction and the preliminary theory concerning random variables that take values in a separable Hilbert space $H$. A general space $L^2(S)$ of such random variables is defined. Examples of the space $H$ are

(i) the real line,

(ii) the complex plane, and

(iii) $R^q$, the euclidean $q$-space.

In Chapter 3 we define a bivariate correlation coefficient, a multiple correlation coefficient and a partial correlation coefficient for $H$-valued random variables. The properties of these coefficients are given and their relationships with the existing theory are discussed.

The partial correlation coefficient is of particular interest since it can be used in two ways;

Received 14 January 1974. Thesis submitted to the Flinders University of South Australia, December 1973. Degree approved, November 1974. Supervisor: Professor J.N. Darroch.



https://doi.org/10.1017/S0004972700023741 Published online by Cambridge University Press

(i) to compare a regression type model and a regression type hypothesis; and

(ii) to measure the improvement in predicting a random variable when more variables or information are added to an existing predictor.

Because the general partial correlation coefficient is the correlation between two particular random variables we can plot these variables to obtain a scatter diagram. The uses of such scatter diagrams are considered and examples given.

In Chapter 4 we use the theory of Chapter 3 to define correlation coefficients for non-negative random variables whose sum is bounded by a constant. These variables are called bounded-sum random variables. Scatter diagrams are given and correlation coefficients are calculated and compared with those obtained by Darroch and Ratcliff [2].

The existing optimal properties of principal components are generalised, in Chapter 5, to $H$-valued random variables and then further generalisations are made. Okamoto [3] shows that the principal components of $X_1, \ldots, X_p$ are optimal, in some senses, over all random variables in $L^2(S)$. By considering only those random variables in $\Omega_0$, a closed linear subspace of $L^2(S)$, we give general results that become those of Okamoto when $\Omega_0 = L^2(S)$. These generalised results are used to answer a conjecture of Rao [4] concerning optimality of principal components of the instrumental variables $Z_1, \ldots, Z_m \in L^2(S)$.

An optimal property of principal component time-series (see Brillinger, [1]) is generalised in Chapter 6. The optimality of the principal component time-series is extended to a larger class of time-series.

In Chapter 7 we consider ways in which particular "linear" combinations of vectors $\mathrm{X}_1, \ldots, \mathrm{X}_p$, taking values in $H^q$, can be defined so that they have optimal properties analogous to those for principal components of $X_1, \ldots, X_p$.

https://doi.org/10.1017/S0004972700023741 Published online by Cambridge University Press

## References

[1] D.R. Brillinger, "The generalisation of the techniques of factor analysis, canonical correlation and principal components to stationary time series", *Roy. Statist. Soc. Conf.* Cardiff, Wales, 1964.

[2] John N. Darroch and Douglas Ratcliff, "Null correlation for proportions - II", *J. Internat. Assoc. Math. Geol.* **2** (1970), 307-312.

[3] Masashi Okamoto, "Optimality of principal components", *Multivariate analysis - II*, 673-685 (Proc. Second Internat. Sympos. Multivariate Analysis, Wright State University, Ohio, 1968. Academic Press, New York and London, 1969).

[4] C.R. Rao, "The use and interpretation of principal component analysis in applied research", *Sankhyā*, *Ser. A* **26** (1964), 329-358.