*Inequalities in Statistics and Probability*
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 165–172

# MOMENT INEQUALITIES WITH APPLICATIONS TO REGRESSION AND TIME SERIES MODELS

BY TZE LEUNG LAI[1] and CHING ZONG WEI[2]
*Columbia University and the University of Maryland*

Herein we review several important moment inequalities in the literature and discuss their applications to strong (almost sure) limit theorems for linear processes and for least squares estimates in multiple regression models.

**1. Introduction and Summary.** A classical model for random noise in the regression and time series literature is that of *equinormed orthogonal* random variables $\epsilon_n$, i.e.,

$$E(\epsilon_i \epsilon_j) = 0 \quad \text{for } i \neq j, \qquad = \sigma^2 \quad \text{for } i = j. \tag{1.1}$$

Such random variables have the important mean square property that for all constants $c_i$,

$$E(\Sigma_{i=m}^n c_i\epsilon_i)^2 = \sigma^2 \Sigma_{i=m}^n c_i^2 \qquad \text{for all } n \geqslant m. \tag{1.2}$$

For example, the so-called Gauss-Markov model in multiple regression theory is of the form

$$z_i = \beta_1 t_{i1} + \ldots + \beta_k t_{ik} + \epsilon_i \, (i=1,2, \ldots ) \tag{1.3}$$

where $t_{ij}$ are known constants, $z_i$ are observed random variables, $\beta_i, \ldots, \beta_k$ are unknown parameters, and $\epsilon_i$ are equinormed orthogonal random variables that represent unobservable random errors. Throughout the sequel we shall let $\mathbf{T}_n$ denote the design matrix $(t_{ij})_{1\leqslant i\leqslant n,\ 1\leqslant j\leqslant k}$, and let $\mathbf{Z}_n = (z_1, \ldots, z_n)'$. For $n\geqslant k$, the least squares estimate $\mathbf{b}_n = (b_{n1}, \ldots, b_{nk})'$ of $\beta = (\beta_1, \ldots, \beta_k)'$ based on the design matrix $\mathbf{T}_n$ and the response vector $\mathbf{Z}_n$ is given by

$$\mathbf{b}_n = (\mathbf{T}'_n\mathbf{T}_n)^{-1}\mathbf{T}'_n\mathbf{Z}_n, \tag{1.4}$$

provided that $\mathbf{T}'_n\mathbf{T}_n$ is nonsingular. From (1.1), it follows easily that

$$\text{cov}(\mathbf{b}_n) = \sigma^2(\mathbf{T}'_n\mathbf{T}_n)^{-1}, \tag{1.5}$$

and therefore $\mathbf{b}_n$ is weakly consistent (i.e., $\mathbf{b}_n \xrightarrow{P} \mathbf{B}$) if

$$(\mathbf{T}'_n\mathbf{T}_n)^{-1} \to \mathbf{0} \text{ as } n\to\infty. \tag{1.6}$$

If $\sigma \neq 0$, the condition (1.6) is also necessary for the weak consistency of $\mathbf{b}_n$ (cf. Drygas (1976)).

In time series theory, it is well known that every wide-sense stationary sequence $\{y_n\}$ with zero means and an absolutely continuous spectral distribution can be represented as

$$y_n = \text{l.i.m.}_{N\to\infty}\Sigma_{i=-N}^N a_{n-i}\epsilon_i, \tag{1.7}$$

where $\{\epsilon_n\}$ is an orthonormal sequence (i.e., $\sigma=1$ in (1.1)), $\{a_n\}$ is a sequence of constants such that $\Sigma_{n=-\infty}^\infty a_n^2 < \infty$, and l.i.m. denotes limit in quadratic mean (cf. Doob (1953), page 499). From this representation, it follows that

[1,2] Research supported by the National Science Foundation.
*AMS 1970 subject classifications:* Primary 60F15; Secondary 60G35, 62J05, 62M10.
Key words and phrases. Least squares theory, linear processes, orthogonal random variables, moment inequalities, lacunary systems, almost sure convergence, law of the iterated logarithm.