Research Results

My research field is statistical science, and my research to present can be broadly classified into the following three categories:

(1) Asymptotic expansion the of null distributions of test statistics in linear models under nonnormality;

(2) Asymptotic behavior of information criteria in model selection problems;

(3) Development of varying coefficient models for spatio-temporal data analysis.

As the results of research belong to the category (1), I have obtained asymptotic expansions for the test statistics of the one-way analysis of variance model, multiple regression model, multivariate linear regression model, and GMANOVA model (Fujikoshi, Ohmae, & Yanagihara, 1999; Yanagihara, 2000, 2001, 2005; Wakaki, Yanagihara, & Fujikoshi, 2002). When actually using asymptotic expansions under nonnormality as approximations for the null distribution, in most cases, it is necessary to estimate the kurtosis of the true distribution. However, it is well known that the usual estimator of kurtosis that is widely used have large biases even with a sufficient sample size. Therefore, I proposed a new estimate of kurtosis that has less bias than existing estimates (Yanagihara, 2007) for improving an asymptotic approximation. For other examples of applications using asymptotic expansions, methods for improving the accuracy of asymptotic approximations by transforming test statistics using monotonic functions based on asymptotic expansion (Yanagihara & Yuan, 2005; Matsumoto, Yanagihara & Wakaki, 2011), and deriving conditions for robustness against nonnormality using asymptotic expansion formulas (Yanagihara, 2007) are also being studied. From this condition, it becomes clear whether a test can be used without worrying about violations of normality.

In the research belong to the category (2), we corrected the bias of information criteria such as AIC (Yanagihara, Sekiguchi & Fujikoshi, 2003; Fujikoshi *et al.*, 2003; Yanagihara *et al.*, 2012; Kamo, Yanagihara & Satoh, 2013). In some linear models, the MC_p criterion, which fully corrects the bias of the C_p criterion, has been obtained (Yanagihara, Nagai & Satoh, 2009; Yanagihara & Satoh, 2010; Yanagihara *et al.*, 2023; Shibayama, Kirishima & Yanagihara, 2024). Furthermore, in a normal multivariate linear model, when the dimension of the response variable vector exceeds the sample size, the AIC cannot be defined because the sample covariance matrix becomes singular. However, by using a ridge-type estimator of the covariance matrix to avoid singularity, we proposed a new AIC that reevaluates the bias term based on a new asymptotic theory where both the dimension and sample size go to infinity simultaneously (Yamamura, Yanagihara & Srivastava, 2010). The above results are for cases where the distribution of the assumed model is the same as that of the true model, but we have also obtained results on bias correction for information criteria in cases where they are different (Yanagihara, 2005; Fujikoshi, Yanagihara & Wakaki, 2005; Yanagihara, 2006; Yanagihara, Himeno & Yuan, 2010; Yanagihara, Kamo & Tonda, 2011; Hashiyama, Yanagihara & Fujikoshi, 2014). Among these, the results of Fujikoshi, Yanagihara and Wakaki (2005) were selected by the members of the Editorial Board to receive the 2006 Jacob Wolfowitz Prize. In addition, we proposed bias-corrected CV criteria under general candidate models (Yanagihara, Tonda & Matsumoto, 2006; Yanagihara & Fujisawa, 2012; Yanagihara *et al.*, 2013). These results are highly generalizable and have simple formulas, making them very broadly applicable.

In the research belong to the category (3), we collaborated with Tokyo Kantei Co., Ltd., a real estate appraisal company, to develop a real estate price prediction model and collaborated with Norway in a bilateral program between Japan and Norway to develop a whale growth prediction model. Real estate price data are highly dependent on the region, such as land prices and the convenience of commercial and public facilities. Therefore, Ohishi *et al.* (2021, 2023, 2025) applied a geographically weighted regression model with a discrete varying coefficient to real estate price data using Fused Lasso, which allows the parameter to be estimated as the same value for adjacent regions with similar estimated values. This method allows for the simultaneous construction of a prediction model and clustering of individuals. We believe that determining the areas where predictions can be made using the same model is an important factor in marketing. Furthermore, this discrete model makes it possible to model the common situation in which trends change dramatically when crossing roads, rivers, or municipal boundaries. For whale growth data, the data are highly time- and location-dependent because food abundance varies with the migration period and location of the whales. Therefore, we developed and analyzed a model that describes the effects of year, seasonal variation, and location on growth using polynomial functions (Yamamura *et al.*, 2016; Solvang *et al.*, 2017).