Journal of Statistical Planning and Inference 171 (2016) 99–114

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Effect of bivariate data's correlation on sequential tests of circular error probability

Yan Li [a], Yajun Mei [b,*]

[a] School of Finance and Statistics, East China Normal University, Shanghai 200241, China

[b] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA

ARTICLE INFO

Article history:
Received 29 January 2015
Received in revised form 28 July 2015
Accepted 12 November 2015
Available online 28 November 2015

MSC:
62F03
62L10

Keywords:
Bivariate normal distribution
Circular error probability (CEP)
Function hypothesis test
Generalized sequential likelihood ratio test (GSLRT)
Sequential probability ratio test (SPRT)

ABSTRACT

The problem of evaluating a military or GPS/GSM system's precision quality is considered in this article, where one sequentially observes bivariate normal data (X_i, Y_i)'s and wants to test hypotheses on the circular error probability (CEP) or the probability of nonconforming, i.e., the probabilities of the system hitting or missing a pre-specified disk target. In such a problem, we first consider a sequential probability ratio test (SPRT) developed under the erroneous assumption of the correlation coefficient $\rho = 0$, and investigate its properties when the true $\rho \neq 0$. It was shown that at least one of the Type I and Type II error probabilities would be larger than the required ones if the true $\rho \neq 0$, and for the detailed effects, $\exp\{-2\} \approx 0.1353$ turns out to be a critical value for the hypothesized probability of nonconforming. Moreover, we propose several sequential tests when the correlation coefficient ρ is unknown, and among these tests, the method of generalized sequential likelihood ratio test (GSLRT) in Bangdiwala (1982) seems to work well.

1. Introduction

Our research is motivated by the problem of evaluating a system's precision quality, which has important applications in the military science of ballistics, GPS (global positioning system) and GSM (global system for mobile communications). In a simple setting of such a problem, one observes two-dimensional vector data $(X_i, Y_i)'$s, and wants to investigate the probabilities of the system (bombs, missiles, bullets, GPS, GSM, etc.) hitting or missing a pre-specified disk target. Mathematically, we want to investigate the circular error probability $CEP = \mathbf{P}(X_i^2+Y_i^2 \le a^2)$, or equivalently, the probability of nonconforming, $p = \mathbf{P}(X_i^2 + Y_i^2 > a^2)$, for some pre-specified constant $a > 0$, and without loss of generality, we assume $a = 1$ hereafter. More formally, the problem of evaluating a system's precision quality can be formulated as testing the hypotheses

$$H_0 : p = p_0 \text{ (“good quality”)} \quad \text{versus} \quad H_1 : p = p_1 \text{ (“poor quality”)}, \tag{1}$$

where $p_0 < p_1$ are pre-specified constants in $(0, 1)$, see Li et al. (2011).

In some real-world applications such as ballistics, it is often very expensive to collect data, and thus it is desired to use as few samples as possible to decide whether the precision quality of the system is good or poor. In the statistical literature, this type of problem is known as sequential hypothesis testing. For the usual statistical tests the sample size is fixed before the

* Corresponding author.
E-mail addresses: yli@stat.ecnu.edu.cn (Y. Li), ymei@isye.gatech.edu (Y. Mei).

http://dx.doi.org/10.1016/j.jspi.2015.11.001
0378-3758/

data are taken, but for a sequential test the total sample size depends on the data and is thus a random variable. Sequential tests are often economical in the sense that we may reach a decision earlier via a sequential test than via a fixed sample size test on average. The sequential test was first investigated by Wald (1945), and since then it has undergone tremendous growth over the past seventy years and has been widely used in applications such as clinical trials and industrial quality control, see Siegmund (1985), Lai (2001) and Tartakovsky et al. (2014) for reviews. In the standard formulation of sequential hypothesis testing, we want to use as few samples as possible to decide which hypothesis in (1) is true, subject to the classical constraints on Type I and Type II error probabilities:

$$\mathbf{P}_{p=p_0}\,(\text{Reject } H_0) \le \alpha \quad \text{and} \quad \mathbf{P}_{p=p_1}(\text{ Reject } H_1) \le \beta, \tag{2}$$

where α and β are pre-specified bounds (usually $0 < \alpha, \beta < 0.5$).

For the sequential hypothesis testing problem in (1) and (2), two existing methods have been proposed in the context of the circular error probability (CEP), see Li et al. (2011). The first one is a naive approach that is based on the Bernoulli data $Z_i = \mathbf{1}(X_i^2 + Y_i^2 > 1)$, where $\mathbf{1}(\cdot)$ is an indicator function. This is a robust nonparametric approach in the sense that it does not make any assumptions on the underlying distributions of two-dimensional vector data (X_i, Y_i)'s, but unfortunately it is rarely used in practice as it needs larger sample sizes on average to make any decisions. The second method, proposed in Li et al. (2011), is a parametric sequential probability ratio test (SPRT) developed under the assumption that the observations $(X_i, Y_i)'$s are independent and identically distributed (i.i.d.) according to a specific bivariate normal distribution $N((0, 0), \ \sigma^2 I_{2\times 2})$, where $I_{2\times 2}$ is a 2×2 identity matrix. However, in a concrete application for the problem in (1) and (2) with $p_0 = 0.1$ and $p_1 = 0.4$, Li et al. (2011) found that their proposed parametric SPRT rejects the null hypothesis $H_0 : p = 0.1$ after only taking $n = 6$ observations even though all the 6 observations (X_i, Y_i)'s fall inside the disk $X_i^2 + Y_i^2 \le 1$. In other words, it declares that the system's precision quality is poor, although the empirical observed probability of nonconforming is $\hat{p} = 0$. Such a disparity has a profound impact in applications to the producers and users of the system, as it can lead completely different conclusions on the precision quality of the system. Assumptions are important in statistical procedures, and thus it is interesting to investigate the effect of the specific $N((0, 0), \ \sigma^2 I_{2\times 2})$ model assumption on the sequential testing problem in (1) and (2).

Note that in the CEP context, it is standard to assume that the (X_i, Y_i)'s are i.i.d. bivariate normal with $\mathbf{E}(X_i) = 0$ and $\mathbf{E}(Y_i) = 0$, as practitioners often adjust the bias of experiment outcomes so as to focus on the precision itself, see Fraser (1951), Fraser (1953), Solomon (1960), Harter (1960), Gillis (1991), Pyati (1993) and Shnidman (1995). Below we took a closer look at those 6 observed data (X_i, Y_i)'s in Li et al. (2011). For this dataset, the point estimates of $\mathbf{E}(X_i)$, $\mathbf{E}(Y_i)$, $Var(X_i)$, $Var(Y_i)$, and $\rho = corr(X_1, Y_i)$ are -0.148, 0.312, 0.326, 0.306 and 0.172, respectively. Hence, the assumptions of equal variances (i.e., $Var(X_i) = Var(Y_i)$) and zero means (i.e., $\mathbf{E}(X_i) = \mathbf{E}(Y_i) = 0$) seem to be reasonable, as $\mathbf{E}(X_i)$ and $\mathbf{E}(Y_i)$ seem to be not significantly different from 0 given the observed variance values. However, the assumption of the correlation coefficient $\rho = corr(X_1, Y_i) = 0$ seems questionable. We took a further step to use Fisher's transformation to find that the 95% confidence interval on ρ is $[-0.65, 0.81]$, but we feel that such a wide confidence interval is due to the small sample size of 6, and does not necessarily prove that ρ is 0. In addition, we should also mention that some weak correlations between X_i's and Y_i's for GPS data were also observed in Amiri-Simkooei (2009) based on the time series datasets from five permanent GPS stations—the correlation coefficients of the (north and east) coordinate components of GPS data are estimated as -0.06, -0.05, 0.07, 0.08 and 0.10 for these five stations, respectively, and Amiri-Simkooei (2009) suspects that these weak correlations were intrinsic for GPS due to the simultaneous estimation of (north and east) coordinate components from the same dataset. Based on the above discussions, it is natural to investigate what happens to the sequential hypothesis testing problem in (1) and (2) under the bivariate normal model with zero means and equal variances when the correlation coefficient ρ between X_i and Y_i is small but not necessarily zero.

The primary objective of this paper is to understand the effect of misusing the correlation coefficient $\rho = 0$ in the sequential hypothesis testing problem in (1) and (2). Intuitively, when the correlation coefficient ρ is not zero, but $\rho = 0$ is erroneously assumed to develop a mis-specified SPRT, called mis-SPRT thereafter, for the problem in (1) and (2), the probabilities of Type I and Type II errors and the average sample numbers of the mis-SPRT will be different from what they should be under $\rho = 0$. Here we will conduct a theoretical analysis to not only justify our intuition, but also gain a deeper understanding how the properties of the mis-SPRT vary as a function of ρ when ρ is small. Our results indicate that when the true correlation coefficient $\rho \ne 0$, the stopping of the mis-SPRT is meaningless in the sense that the observed data still cannot provide enough evidence to decide which hypothesis in (1) is true subject to the Type I and Type II error probability constraints in (2).

A closely related objective of this paper is to develop a sequential test when the correlation coefficient ρ is unknown. While the sequential hypothesis testing problem with unknown nuisance parameters have been well studied in the literature, see, Wald (1947), to the best of our knowledge, no sequential tests have been developed in the context of CEP for two-dimensional random vector data. The main challenge is the computational complexity in the context of CEP when the correlation coefficient $\rho \ne 0$. Here we demonstrate how to modify several existing sequential methods to the CEP context, and hopefully it will stimulate further research on sequential tests for CEP.

We should mention that the case of model or distributional mis-specification has been investigated in the statistical literature, though most, if not all, on the offline fixed-sample setting. For instance, Foutz and Srivastava (1977) considers the property of the fixed-sample likelihood ratio test when the (normal distribution) model is mis-specified. On the other hand, while much research has been done on sequential tests of the bivariate normal model, or more generally, d-dimensional