Genetic association tests when a nuisance parameter is not identifiable under no association

Wonkuk Kim [1,a], Yeong-Hwa Kim [a]

[a]Department of Applied Statistics, Chung-Ang University, Korea

Abstract

Some genetic association tests include an unidentifiable nuisance parameter under the null hypothesis of no association. When the mode of inheritance (MOI) is not specified in a case-control design, the Cochran-Armitage (CA) trend test contains an unidentifiable nuisance parameter. The transmission disequilibrium test (TDT) in a family-based association study that includes the unaffected also contains an unidentifiable nuisance parameter. The hypothesis tests that include an unidentifiable nuisance parameter are typically performed by taking a supremum of the CA tests or TDT over reasonable values of the parameter. The p-values of the supremum test statistics cannot be obtained by a normal or chi-square distribution. A common method is to use a Davies's upper bound of the p-value instead of an exact asymptotic p-value. In this paper, we provide a unified sine-cosine process expression of the CA trend test that does not specify the MOI and the TDT that includes the unaffected. We also present a closed form expression of the exact asymptotic formulas to calculate the p-values of the supremum tests when the score function can be written as a linear form in an unidentifiable parameter. We illustrate how to use the derived formulas using a pharmacogenetics case-control dataset and an attention deficit hyperactivity disorder family-based example.

Keywords: Cochran-Armitage trend test, Rice's formula, sine-cosine process, transmission disequilibrium test, unidentifiable parameter

1. Introduction

A genome-wide association study is a powerful method to screen a high-dimensional genome data set and select candidate single nucleotide polymorphisms (SNPs) for genetic associations. Genetic association studies are commonly conducted by a case-control design or family-based design. Table 1 shows a summarization of a case-control data set at a single biallelic SNP. The Cochran-Armitage (CA) linear trend test (Armitage, 1955; Cochran, 1954) is known to be a powerful test in a genetic case-control association study, but it requires the specification of a mode of inheritance (MOI). The MOI can be specified by selecting a weight vector $(0, \theta, 1)$ where $\theta = 0$ for recessive model, $\theta = 1/2$ for additive model, and $\theta = 1$ for dominant model. The CA linear trend test can be written as

$$Z_{\text{CA}}(\theta) = \frac{\sqrt{n}[\theta(rs_1 - sr_1) + (rs_2 - sr_2)]}{\sqrt{rs\left[\theta^2 n_1(n - n_1) - 2n_1 n_2 \theta + n_2(n - n_2)\right]}}. \tag{1.1}$$

The CA trend test statistic is a Rao's score test statistic that can be derived from the logit model of

$$\log\left(\frac{P(Y = 1|G)}{1 - P(Y = 1|G)}\right) = \alpha + \beta\left\{\theta \cdot I(G = 1) + I(G = 2)\right\}, \tag{1.2}$$

---

[1] Corresponding author: Department of Applied Statistics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea. E-mail: wkim@cau.ac.kr

Table 1: Data structure in a case-control association study

	AA	AB	BB	Total
Control	r_0	r_1	r_2	r
Case	s_0	s_1	s_2	s
Total	n_0	n_1	n_2	n

Table 2: Data structure in a family-based association study

	Transmitted	Not transmitted	Total
Unaffected	b_U	c_U	n_0
Affected	b_A	c_A	n_1

where $I(\cdot)$ is an indicator function and G is the number of allele B in Table 1, that is, $G = 0$ for genotype AA, $G = 1$ for genotype AB and $G = 2$ for genotype BB. Under the null hypothesis of no association, that is, $H_0 : \beta = 0$, the nuisance parameter θ is not estimable and hence it is not identifiable under the null hypothesis.

The family-based association study is known to be less powerful than the population-based association study, but the transmission disequilibrium test (TDT) (Spielman *et al.*, 1993) may be a unique robust test that does not depend on the phenotype distribution, genotype frequencies, or mating distribution. The power of the TDT may increase by including the unaffected in the statistical analysis (Lange and Laird, 2002; Lunetta *et al.*, 2000). Table 2 summarizes a single-locus data set in a family-based association study. The TDT including the unaffected is given by

$$Z_{\text{TDT}} = \frac{(1 - \mu)(b_A - c_A) - \mu(b_U - c_U)}{\sqrt{(1 - \mu)^2(b_A + c_A) + \mu^2(b_U + c_U)}}, \tag{1.3}$$

where the nuisance parameter μ is not identifiable under the null hypothesis of no association. One of ? common choices of the parameter μ is the prevalence of disease, that is not usually estimable in the case-control or family-based design.

Davies (1977) proposed the supremum test statistics when a nuisance parameter is present only under the alternative hypothesis. In addition, he presented a general formula for an upper bound of the p-value of the supremum test. This upper bound in an integral form is generally called the Davies' upper bound.

Gaussian processes are widely used in many scientific fields (Choi and Lee, 2014; Lee and Park, 2017). A sine-cosine process is the simplest Gaussian wave process written as

$$U(t) = U_1 \cos t + U_2 \sin t \tag{1.4}$$

where U_1 and U_2 are independent standard normal random variables. Delmas (2003) derived the exact formula to calculate the upper tail probability of $M(t) = \sup_{0 \le s \le t} U(s)$ by applying Rice's formula (Rice, 1944, 1945).

In this work, we show that the supremum test of the CA linear trend test and the supremum test of the TDT can be written in terms of the supremum of a sine-cosine process, and we provide the exact asymptotic p-value calculation formulas for unified test statistics based on the supremum of the sine-cosine process.

2. Methods

Suppose that the probability function $p(y, \beta, \alpha, \phi)$ does not depend on θ when the null hypothesis is true, that is, $H_0 : \beta = \beta_0$. In this work, we assume β and θ are scalars and α can be a parameter vector.

The score function is written as

$$U_{\beta,N}(\beta_0, \alpha, \phi) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\partial}{\partial \beta}\bigg|_{\beta=\beta_0} \log p(y_i, \beta, \alpha, \phi). \tag{2.1}$$

Let $\hat{\alpha}_0$ be the maximum likelihood estimate of α under $H_0 : \beta = \beta_0$. Now suppose this score function can be written in form of

$$U_{\beta,N}(\beta_0, \hat{\alpha}_0, \phi) = X_{1,N} + w(\phi)X_{2,N} \tag{2.2}$$

where w is a continuous function in $\mathbb{R}$ and let $\theta = w(\phi)$. Suppose that the bivariate random vector $X_N = (X_{1,N}, X_{2,N})'$ converges in distribution to a mean zero bivariate normal random vector, that is, $X_N = (X_{1,N}, X_{2,N})' \xrightarrow{d} X = (X_1, X_2)' \sim N((0,0)', \Sigma)$ where $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ as $N \to \infty$ under the null hypothesis. For simplicity, we use notation

$$U_{\beta,N}(\beta_0, \hat{\alpha}_0, \theta) = X_{1,N} + \theta X_{2,N} \tag{2.3}$$

and we define the locally most powerful test given θ as

$$Z_N(\theta) = \frac{U_{\beta,N}(\beta_0, \hat{\alpha}_0, \theta)}{\sqrt{\sigma_1^2 + 2\theta\rho\sigma_1\sigma_2 + \theta^2\sigma_2^2}} = \frac{X_{1,N} + \theta X_{2,N}}{\sqrt{\sigma_1^2 + 2\theta\rho\sigma_1\sigma_2 + \theta^2\sigma_2^2}}. \tag{2.4}$$

Since $Z_N(\theta) \xrightarrow{d} Z(\theta)$, the p-value of the test statistic can be obtained from the Gaussian process of

$$Z(\theta) = \frac{\left[X_1 - \frac{\sigma_1}{\sigma_2}\rho X_2\right] + \left(\theta + \frac{\sigma_1}{\sigma_2}\rho\right) X_2}{\sqrt{\sigma_1^2 + 2\theta\rho\sigma_1\sigma_2 + \theta^2\sigma_2^2}}. \tag{2.5}$$

We define two independent standard normal random variable U_1 and U_2 as

$$U_1 = \frac{X_1 - \frac{\sigma_1}{\sigma_2}\rho X_2}{\sigma_1\sqrt{1-\rho^2}}, \qquad U_2 = \frac{X_2}{\sigma_2}. \tag{2.6}$$

Using these standard random variables, we can write

$$Z(\theta) = \frac{\sigma_1\sqrt{1-\rho^2}U_1 + (\rho\sigma_1 + \theta\sigma_2)U_2}{\sqrt{\left(\sigma_1\sqrt{1-\rho^2}\right)^2 + (\rho\sigma_1 + \theta\sigma_2)^2}}. \tag{2.7}$$

We define an angle $t = t(\theta)$ by

$$\cos t = \frac{\sigma_1\sqrt{1-\rho^2}}{\sqrt{\left(\sigma_1\sqrt{1-\rho^2}\right)^2 + (\rho\sigma_1 + \theta\sigma_2)^2}}, \qquad \sin t = \frac{\rho\sigma_1 + \theta\sigma_2}{\sqrt{\left(\sigma_1\sqrt{1-\rho^2}\right)^2 + (\rho\sigma_1 + \theta\sigma_2)^2}}. \tag{2.8}$$

Here $\cos t$ is always nonnegative for all θ but $\sin t$ can be positive or negative depending upon θ. Hence we may assume that $-\pi/2 \le t \le \pi/2$. In terms of the parameter $t = t(\theta)$,

$$Z(\theta) = U_1 \cos t(\theta) + U_2 \sin t(\theta). \tag{2.9}$$

Let $U(t) = U_1 \cos t + U_2 \sin t$ where $t(\theta) = \arctan((\rho\sigma_1 + \theta\sigma_2)/(\sigma_1 \sqrt{1-\rho^2}))$. Let $t_L = \min_{\theta_L \leq \theta \leq \theta_U} t(\theta)$ and $t_U = \max_{\theta_L \leq \theta \leq \theta_U} t(\theta)$. Here $U(t)$ is the orthogonal projection of (U_1, U_2) onto the straight line joining the origin $(0, 0)$ and $(\cos t, \sin t)$. Without loss of generality, suppose $t_U - t_L \leq \pi$ since $-\pi/2 \leq t \leq \pi/2$. First we can simply obtain the supremum of $U(t)$ or $|U(t)|$ by the following equations:

$$\sup_{t_L \leq t \leq t_U} U(t) = \begin{cases} \sqrt{U_1^2 + U_2^2}, & \text{for } t_L \leq \arctan\left(\frac{U_2}{U_1}\right) \leq t_U, \ U_1 \geq 0, \\ \max_{t \in \{t_L, t_U\}} (U_1 \cos t + U_2 \sin t), & \text{otherwise,} \end{cases} \quad (2.10)$$

$$\sup_{t_L \leq t \leq t_U} |U(t)| = \begin{cases} \sqrt{U_1^2 + U_2^2}, & \text{for } t_L \leq \arctan\left(\frac{U_2}{U_1}\right) \leq t_U, \\ \max_{t \in \{t_L, t_U\}} |U_1 \cos t + U_2 \sin t|, & \text{otherwise.} \end{cases} \quad (2.11)$$

For a given $u > 0$, we can calculate the right-tail probabilities of $\sup U(t)$ and $\sup |U(t)|$ by the following Theorem.

Theorem 1. *Suppose $-\pi/2 \leq t_L \leq t_U \leq \pi/2$ and $u > 0$. Let $U(t)$ be a sine-cosine process defined on $[t_L, t_U]$. Let $\Phi(u)$ be the distribution function of a standard normal random variate.*

1. The one-sided p-value is the same as the Davies' upper bound that is given by

$$P\left(\sup_{t_L \leq t \leq t_U} U(t) \geq u\right) = 1 - \Phi(u) + \frac{t_U - t_L}{2\pi} e^{-\frac{u^2}{2}}. \quad (2.12)$$

2. The two-sided p-value is smaller than the Davies' upper bound and it can be calculated by

$$P\left(\sup_{t_L \leq t \leq t_U} |U(t)| \geq u\right) = 2\left[1 - \Phi(u)\right] + \frac{t_U - t_L}{\pi} e^{-\frac{u^2}{2}} - \frac{1}{\pi} \int_0^{t_U - t_L} \exp\left[-\frac{u^2}{1 - \cos s}\right] ds \quad (2.13)$$

$$= 2P\left(\sup_{t_L \leq t \leq t_U} U(t) \geq u\right) - \frac{1}{\pi} \int_0^{t_U - t_L} \exp\left[-\frac{u^2}{1 - \cos s}\right] ds. \quad (2.14)$$

If $t_U - t_L = \pi/2$, then

$$P\left(\sup_{t_L \leq t \leq t_U} |U(t)| \geq u\right) = 2\left[1 - \Phi(u)\right] + \frac{1}{2} e^{-\frac{u^2}{2}} - 2\left[1 - \Phi(u)\right]^2. \quad (2.15)$$

The proof of Theorem 1 is given in Appendix. The following corollary is immediate from the contents in this section and Theorem 1.

Corollary 1. (Unified representation of $Z_{\text{TDT}}(\mu)$ and $Z_{\text{CA}}(\theta)$) $Z_{TDT}(\mu)$ and $Z_{CA}(\theta)$ can be written as a sine-cosine process so that the supremum tests are

$$\sup_{t_l \leq t \leq t_u} U(t) \quad or \quad \sup_{t_l \leq t \leq t_u} |U(t)|, \quad (2.16)$$

where $U(t) = U_1 \cos t + U_2 \sin t$ is a sine-cosine process with independent standard normal random variates U_1 and U_2. For the family-based association test,

$$\cos t = \frac{(1 - \mu) \sqrt{b_A + c_A}}{\sqrt{(1 - \mu)^2 (b_A + c_A) + \mu^2 (b_U + c_U)}}, \qquad \sin t = \frac{\mu \sqrt{b_U + c_U}}{\sqrt{(1 - \mu)^2 (b_A + c_A) + \mu^2 (b_U + c_U)}}. \quad (2.17)$$

If $0 \leq \mu_L \leq \mu \leq \mu_U \leq 1$, then

$$t_L = \arctan\left(\frac{\mu_L}{1-\mu_L}\sqrt{\frac{b_U + c_U}{b_A + c_A}}\right) \quad and \quad t_U = \arctan\left(\frac{\mu_U}{1-\mu_U}\sqrt{\frac{b_U + c_U}{b_A + c_A}}\right). \tag{2.18}$$

For the CA linear trend test,

$$\cos t = \frac{\sqrt{1-\rho^2}\sigma_1}{\sqrt{(1-\rho^2)\sigma_1^2 + (\sigma_1\rho + \theta\sigma_2)^2}}, \qquad \sin t = \frac{\sigma_1\rho + \theta\sigma_2}{\sqrt{(1-\rho^2)\sigma_1^2 + (\sigma_1\rho + \theta\sigma_2)^2}}, \tag{2.19}$$

where $\sigma_1 = \sqrt{p_2(1-p_2)}$, $\sigma_2 = \sqrt{p_0(1-p_0)}$, and $\rho = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}}$,

$$t_L = -\arctan\left(\sqrt{\frac{p_1 p_2}{p_0}}\right), \qquad t_U = \arctan\left(\sqrt{\frac{p_0 p_1}{p_2}}\right). \tag{2.20}$$

Here, $p_0 = P(AA)$, $p_1 = P(AB)$, and $p_2 = P(BB)$. The p-values of the supremum tests of the TDT including the unaffected and the CA trend test with unspecified MOI can be obtained from Theorem 1.

3. Examples

In this section, we illustrate two examples. One is a case-control type pharmacogenetics data set of anti-epileptic drug responses. The other example is the case-parent trio dataset of attention deficit hyperactivity disorder (ADHD) illustrated in Lunetta *et al.* (2000).

3.1. Case-control association study

Two hundred and eighty-eight patients of epilepsy were recruited from multiple epilepsy clinics in Korea and they were genotyped for whole-exomes by the next-generation sequencing experiments. All study participants were eligible if they had drug-resistant (case group) or drug-responsive (control group) epilepsy according to the following definitions and criteria explained in Kim *et al.* (2011). Drug resistance was defined as the occurrence of at least four unprovoked seizures in the course of the year before recruitment, with trials of two or more appropriate antiepileptic drugs (AEDs) at maximal tolerated doses. Patients who underwent surgical treatment for drug-resistant epilepsy were classified as having drug-resistant epilepsy, regardless of the surgical outcome. Patients who were frequently in poor compliance with AED therapy and those who had reported seizures with a questionable semiology were excluded from this study. We define drug responsiveness as complete freedom from seizures for at least one year up to the date of the last follow-up visit.

We performed the CA trend tests for recessive, additive, and dominant genetic models and the supremum test of the CA test for undetermined MOI but $0 \leq \theta \leq 1$. The *p*-values of the CA trend tests for the three specific genetic models are calculated by the upper tails of a chi-squared distribution with one degree of freedom while the *p*-values of the supremum tests are calculated by two ways. Table 3 shows the five smallest *p*-values of the supremum of CA trend tests for $0 \leq \theta \leq 1$ calculated by Equation (2.14) in Theorem 1 and the permutation *p*-values from 1 million and 10 millions permuted data. The *p*-values of the CA trend tests under recessive ($\theta = 0$), additive ($\theta = 0.5$), and dominant ($\theta = 1$) MOI are included in the table. For $0 \leq \theta \leq 1$, the permutation *p*-value tends to be smaller than the exact asymptotic *p*-value from Equation (2.14). It appears that the permuted *p*-values based on 10 millions permuted data are closer to the exact asymptotic *p*-values derived in this paper than the permuted *p*-values obtained from one million permuted data.

Table 3: The *p*-values of the top five SNPs selected by the CA trend tests for $0 \le \theta \le 1$ of the anti-epileptic drug data

SNP	$0 \le \theta \le 1$			$\theta = 0$	$\theta = 0.5$	$\theta = 1$
	Exact	1,000,000	10,000,000			
rs16964316	5.42×10^{-6}	1.00×10^{-6}	2.60×10^{-6}	4.68×10^{-6}	2.17×10^{-6}	9.55×10^{-5}
rs17671352	1.34×10^{-5}	7.00×10^{-6}	8.70×10^{-6}	1.42×10^{-1}	1.66×10^{-4}	2.02×10^{-6}
rs16909651	6.09×10^{-5}	5.60×10^{-5}	5.92×10^{-5}	3.14×10^{-5}	2.14×10^{-5}	1.61×10^{-3}
rs12417255	6.33×10^{-5}	5.50×10^{-5}	6.24×10^{-5}	2.75×10^{-3}	1.96×10^{-5}	2.22×10^{-5}
rs12041477	8.84×10^{-5}	7.41×10^{-5}	7.57×10^{-5}	3.25×10^{-2}	7.54×10^{-5}	1.61×10^{-5}

The numbers in the second column are obtained from Equation (2.14). The numbers in the third and fourth columns of $0 \le \theta \le 1$ are the *p*-values based on the Monte-Carlo permutation method from 1 million and 10 millions permuted data, respectively. SNP = single nucleotide polymorphism.

Table 4: Two-tailed *p*-values when the nuisance parameter μ is known to be in an interval, (0, 1), (0.05, 0.1), or (0.114, 0.161) for the attention deficit hyperactivity disorder data

Allele	N				μ	Asymptotic *p*-value	Permutation *p*-value
	b_A	c_A	b_U	c_U			
DAT-480	17	10	6	13	(0.000, 1.000)	0.09099	0.083
					(0.050, 0.100)	0.14124	0.146
					(0.114, 0.161)	0.11762	0.146
DRD4-7	15	6	5	10	(0.000, 1.000)	0.05017	0.058
					(0.050, 0.100)	0.03970	0.040
					(0.114, 0.161)	0.03360	0.034

The Monte Carlo permutation *p*-values are obtained based on 1,000 permuted data.

3.2. Family-based association study

Lunetta *et al.* (2000) recruited 39 nuclear families in which at least one family member was believed to have ADHD and all individuals were assessed for the Diagnostic and Statistical Manual of Mental Disorders IV ADHD to determine their phenotypes. For DAT gene, 60 parent-child trios in 35 nuclear families, consisting of 33 affected and 27 unaffected children, were genotyped. A total of 15 unaffected children had exactly one heterozygous parent; 2 unaffected children had two heterozygous parents. A total of 15 affected children had one heterozygous parent; 6 affected children had two heterozygous parents. Twelve additional trios were genotyped for DRD4 gene, resulting in a total of 72 trios, with 42 affected and 30 unaffected children, in 39 nuclear families. For DRD4-7, 11 unaffected children had exactly one heterozygous parent; and 2 unaffected children had two heterozygous parents. A total of 13 affected children had exactly one heterozygous parent; 4 affected children had two heterozygous parents. The two genes were treated as biallelic; DRD4-7 against all other DRD4 alleles and DAT-480 against all other DAT alleles. For an illustration of the TDT example, we performed the TDT under no information of the prevalence of ADHD, that is, $0 \le \mu \le 1$. We also conducted the TDT under $0.05 \le \mu \le 0.1$ (Lunetta *et al.*, 2000) and under $0.114 \le \mu \le 0.161$ (Faraone *et al.*, 2003). Table 4 shows the *p*-values of the TDTs under two cases. The *p*-value under $0.05 \le \mu \le 0.1$ is greater than the *p*-value under $0 \le \mu \le 1$ for DAT-480 whereas the opposite holds for DRD4-7.

4. Discussion and conclusions

In this paper, we derived simple formulas to calculate the *p*-values of the supremum tests when a score function is linear in an unidentifiable nuisance parameter as in Equation (2.2). The derived formulas can be used to calculate the *p*-values of the CA trend test when the MOI is not specified as

Table 5: The results of the simulation study

SNP used for simulated data	$0 \le \theta \le 1$		
	Exact asymptotic	1,000,000	10,000,000
rs16964316	4.39×10^{-8}	0.00	1.00×10^{-7}
rs17671352	3.12×10^{-6}	2.00×10^{-6}	2.50×10^{-6}
rs16909651	5.05×10^{-7}	0.00	4.00×10^{-7}
rs12417255	9.66×10^{-7}	1.00×10^{-6}	9.00×10^{-7}
rs12041477	2.23×10^{-6}	2.00×10^{-6}	2.70×10^{-6}

Each dataset was generated by three copies of the genotypes of one of SNPs in Table 3. The phenotypes were generated by combining two copies of the phenotypes of the real data and a randomized copy of the phenotypes.
SNP = single nucleotide polymorphism.

well as the TDT including the unaffected. In particular, the Davies's upper bound is the same as the exact asymptotic p-value of the supremum test for a one-sided alternative hypothesis while the exact asymptotic two-tailed p-value is smaller than the corresponding Davies's upper bound. As shown in Examples section, the approximate p-values of the supremum tests can be obtained by the Monte-Carlo permutation method. The exact asymptotic p-value tends to be greater than the permutation p-value in the case-control example, when the sample size is 288. To investigate the convergence of the exact asymptotic p-values, we conducted a small simulation study, in which we generated five fictitious SNPs data. For each SNP in Table 3, we used a dataset having three copies of genotypes while two copies of the phenotypes are used and a copy of phenotypes is randomized. Hence, the sample size of the simulation study is set as 864. By doing this, we could simulate the p-values around 10^{-6}–10^{-8}.

Table 5 shows the simulation results. We could not see any pattern that one method provides smaller p-values compared to the other method. It appears that the permutation p-value may be preferred when the sample size is not sufficiently large, whereas the exact asymptotic p-value may be preferred when the sample size is large enough. However, this permutation approach is computationally intensive and the Monte-Carlo permutation p-values depend on specified seeds and the number of resamples; 175 minutes on average were required to complete an empirical p-value calculation based on 10 million permuted data using a 3.5 GHz intel Xeon processor. Therefore, the permutation method for whole-exome or whole-genome data may not be feasible. Kim *et al.* (2012) and Kim (2015) proposed the linear trend tests and the TDT based on read counts for low-coverage next-generation sequences experiments in which genotypes are uncertain. Their extended tests require much computational resources due to estimating the parameters of mixture models. Our work in this paper can be applied to the extended linear trend tests and the TDT based on read counts while saving computational resources.

We also provided the unified sine-cosine process expression of the supremum tests for the CA linear trend without specifying the MOI and the TDT including the unaffected. This work focused on the case in which there is a one-dimensional unidentifiable nuisance parameter in a linear form. Davies (1987, 2002) considered a chi-square process and an F-process for more complex models and our work may be extended to cases in which two or more unidentifiable nuisance parameters exist.

Appendix:

Proof of Theorem 1: Since the distribution of (U_1, U_2) is invariant under rotations, we may assume $t_L = 0$ and $t_U = s$.

1. For a given $u > 0$,

$$
\begin{aligned}
P&\left(\sup_{0 \leq t \leq s} U(t) \geq u\right) \\
&= \int_{-\frac{\pi}{2}}^{0} \left[\int_{u/\cos\theta}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta + \int_{s}^{s+\frac{\pi}{2}} \left[\int_{u/\cos\theta}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta + \int_{0}^{s} \left[\int_{u}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\int_{u/\cos\theta}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta + \int_{0}^{s} \left[\int_{u}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta \\
&= \int_{u}^{\infty} \left[\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dy\right] dx + \int_{0}^{s} \left[\int_{u}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta \\
&= 1 - \Phi(u) + \frac{s}{2\pi} e^{-\frac{u^2}{2}}.
\end{aligned}
$$

2. For a fixed $u > 0$,

$$
\begin{aligned}
P\left(\sup_{0 \leq t \leq s} |U(t)| \geq u\right) &= 2P\left(\sup_{0 \leq t \leq s} U(t) \geq u\right) - 2 \int_{0}^{\frac{s}{2}} \left[\int_{\frac{u}{\sin\theta}}^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r\,dr\right] d\theta \\
&= 2\left[1 - \Phi(u)\right] + \frac{s}{\pi} e^{-\frac{u^2}{2}} - \frac{1}{\pi} \int_{0}^{s} \exp\left[-\frac{u^2}{1-\cos\theta}\right] d\theta.
\end{aligned}
$$

If $s = \pi/2$, then

$$
\begin{aligned}
P\left(\sup_{0 \leq t \leq \frac{\pi}{2}} |U(t)| \geq u\right) &= 2\left[1 - \Phi(u)\right] + \frac{1}{2} e^{-\frac{u^2}{2}} - 2 \int_{u}^{\infty} \left[\int_{u}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dx\right] dy \\
&= 2\left[1 - \Phi(u)\right] + \frac{1}{2} e^{-\frac{u^2}{2}} - 2\left[1 - \Phi(u)\right]^2.
\end{aligned}
$$

$\square$

Acknowledgements

References

Armitage P (1955). Tests for linear trends in proportions and frequencies, *Biometrics*, **11**, 375–386.

Choi ML and Lee J (2014). GLR charts for simultaneously monitoring a sustained shift and a linear drift in the process mean, *Communications for Statistical Applications and Methods*, **21**, 69–80.

Cochran WG (1954). Some methods for strengthening the common χ^2 tests, *Biometrics*, **10**, 417–451.

Davies RB (1977). Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, **64**, 247–254.

Davies RB (1987). Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, **74**, 33–43.

Davies RB (2002). Hypothesis testing when a nuisance parameter is present only under the alternative: Linear model case, *Biometrika*, **89**, 484–489.

Delmas C (2003). Projections on spherical cones, maximum of Gaussian fields and Rice's method, *Statistics & Probability Letters*, **64**, 263–270.

Faraone SV, Sergent J, Gillberg C, and Biederman J (2003). The worldwide prevalence of ADHD: is it an American condition? *World Psychiatry*, **2**, 104–113.

Kim MK, Moore JH, Kim JK, *et al.* (2011). Evidence for epistatic interactions in antiepileptic drug resistance, *Journal of Human Genetics*, **56**, 71–76.

Kim W (2015). Transmission disequilibrium tests based on read counts for low-coverage next-generation sequence data, *Human Heredity*, **80**, 36–49.

Kim W, Londono D, Zhou L, *et al.* (2012). Single-variant and multi-variant trend tests for genetic association with next-generation sequencing that are robust to sequencing error, *Human Heredity*, **74**, 172–183.

Lange C and Laird NM (2002). Power calculations for a general class of family-based association tests: dichotomous traits, *The American Journal of Human Genetics*, **71**, 575–584.

Lee Y and Park JS (2017). Model selection algorithm in Gaussian process regression for computer experiments, *Communications for Statistical Applications and Methods*, **24**, 383–396.

Lunetta KL, Faraone SV, Biederman J, and Laird NM (2000). Family-based tests of association and linkage that use unaffected sibs, covariates, and interactions. *The American Journal of Human Genetics*, **66**, 605–614.

Rice SO (1944). Mathematical analysis of random noise, *Bell Labs Technical Journal*, **23**, 282–332.

Rice SO (1945). Mathematical analysis of random noise, *The Bell System Technical Journal*, **24**, 46–156.

Spielman RS, McGinnis RE, and Ewens WJ (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM), *The American Journal of Human Genetics*, **52**, 506–516.