UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Friday, 4 June, 2021 12:00 pm to 2:00 pm

PAPER 225

FUNCTIONAL DATA ANALYSIS

Before you begin please read these instructions carefully

Candidates have TWO HOURS to complete the written examination.

Attempt no more than ***THREE*** *questions.*
There are ***FOUR*** *questions in total.*
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

Let X be a Gaussian random element on $L^2[0,1]$, such that $\mathbb{E}\|X\|^2 < \infty$, $\mathbb{E}X = 0$, and suppose the covariance operator C_X of X has covariance kernel

$$c_X(t,s) = \min(t,s) - ts, \quad t,s \in [0,1].$$

(i) Find the eigenfunctions, ϕ_k, and eigenvalues, λ_k, of C_X.

(ii) Assume we have a Gaussian random element X_1 with the same distribution as X, and let

$$X_1(t) = \sum_{k=1}^{\infty} a_{1k}\phi_k(t), \quad t \in [0,1]$$

where a_{1k} are univariate random variables. If we have observations of X_1 at times $t_1, t_2 \in [0,1]$, find the conditional expectation

$$\mathbb{E}(a_{1k}|X_1(t_1), X_1(t_2))$$

for a fixed k.

[You may use the conditional normal distribution formula without proof.]

2 Let $X, X_1, \ldots, X_n$ be i.i.d. elements of $L^2[0,1]$ with associated norm $\|\cdot\|$ such that $\mathbb{E}\|X\|^4 < \infty$, $\mathbb{E}X = \mu$, and with covariance operator C_X. Also let $X^*, X_1^*, \ldots, X_n^*$ be i.i.d. elements of $L^2[0,1]$ such that $\mathbb{E}\|X^*\|^4 < \infty$, $\mathbb{E}X^* = \mu^*$ and with covariance operator C_{X^*}. Also assume the two samples are independent of one another.

By considering the eigendecomposition of $\frac{1}{2}(C_X + C_{X^*})$, for some $K \in \mathbb{N}$ find a K-dimensional test to determine whether $\mu = \mu^*$. Determine its asymptotic properties, as $n \to \infty$, under $H_0 : \mu = \mu^*$ and under the alternative $H_A : \mu \neq \mu^*$.

[You may assume that the first K eigenvalues of $(C_X + C_{X^*})$ are all distinct, and you may also use without proof the convergence properties of eigenvalues, eigenfunctions and any version of the central limit theorem].

3 Let $C_1, \ldots, C_n$ be covariance operators on a separable Hilbert Space.

(i) Define the square-root distance, d_R, between two covariance operators and find

$$\widehat{C}_R = \underset{C \in \mathfrak{C}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} d_R^2(C_i, C) \right\},$$

where $\mathfrak{C}$ is the space of covariance operators on the Hilbert space.

(ii) Suppose $C_1 = L_1 L_1^*$ and $C_2 = L_2 L_2^*$ where L_i are operators on the Hilbert Space, and where L_i^* is the adjoint operator of L_i. Show that $L_2^* L_1$ is trace class.

(iii) Define the Procrustes distance, d_P, between two covariance operators and show that

$$d_P^2(C_1, C_2) = \|L_1\|_{HS}^2 + \|L_2\|_{HS}^2 - 2 \sum_{k=1}^{\infty} \sigma_k,$$

where σ_k are the singular values of $L_2^* L_1$, and where $\| \cdot \|_{HS}$ is the Hilbert-Schmidt norm.

(iv) Let $C_1(\cdot) = \sum_{k=1}^{\infty} \lambda_k \langle \cdot, \phi_k \rangle \phi_k$ and $C_2(\cdot) = \sum_{k=1}^{\infty} \lambda_k^* \langle \cdot, \phi_k \rangle \phi_k$ where ϕ_k is an orthonormal basis of $L^2[0,1]$ and $\lambda_k, \lambda_k^* \in \mathbb{R}^+$, such that $\sum_{k=1}^{\infty} \lambda_k < \infty$ and $\sum_{k=1}^{\infty} \lambda_k^* < \infty$.

Find the square-root and Procrustes distances between C_1 and C_2 and briefly comment on the relationship between the two distances in this case.

4 Let X, ϵ be independent random elements in $L^2[0,1]$, such that $\mathbb{E}\|X\|^2 < \infty$, $\mathbb{E}X = 0$ and $\mathbb{E}\|\epsilon\|^2 < \infty$, $\mathbb{E}\epsilon = 0$, and define

$$Y(t) = \int_0^1 \beta(t,s)X(s)ds + \epsilon(t), \quad t \in [0,1],$$

where $\beta(t,s) \in L^2([0,1] \times [0,1])$.

(i) Suppose that the covariance operator of X is positive definite with eigenvalues / eigenfunctions $\{\lambda_k, \phi_k\}_{k=1}^{\infty}$, and that the covariance operator of Y has eigenvalues / eigenfunctions $\{\gamma_l, u_l\}_{l=1}^{\infty}$. Also suppose that $X = \sum_{k=1}^{\infty} a_k \phi_k$ and $Y = \sum_{l=1}^{\infty} b_l u_l$ where a_k and b_l are univariate random variables. Show that

$$\beta(t,s) = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{\mathbb{E}(a_k b_l)}{\lambda_k} \phi_k(s) u_l(t), \quad t, s \in [0,1].$$

(ii) Show that the above expansion of $\beta(t,s)$ does not change with the choice of sign of the eigenfunctions ϕ_k and u_l.

(iii) Let

$$R(t) = \frac{\text{Var}(\mathbb{E}(Y(t)|X))}{\text{Var}(Y(t))}, \quad t \in [0,1].$$

Find an expression for $R(t)$ in terms of the expansions of Y and X from part (i) and show that this also does not depend on the choice of sign of the eigenfunctions.

END OF PAPER