ONE-SAMPLE PROBLEM FOR FUNCTIONAL DATA

2015 03 17

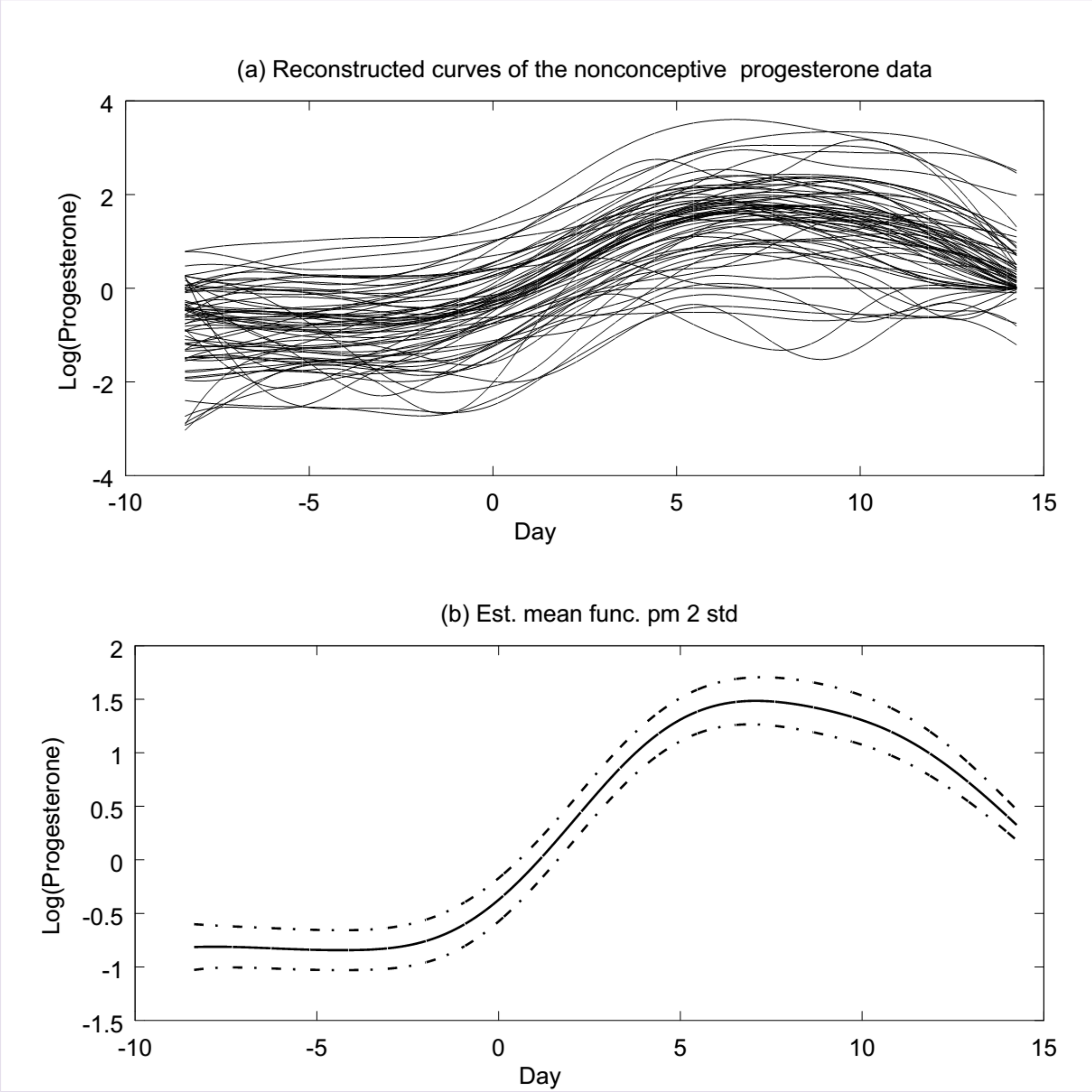
FDA DARBINIS SEMINARAS

INTRODUCTION AND GENERAL SETTING

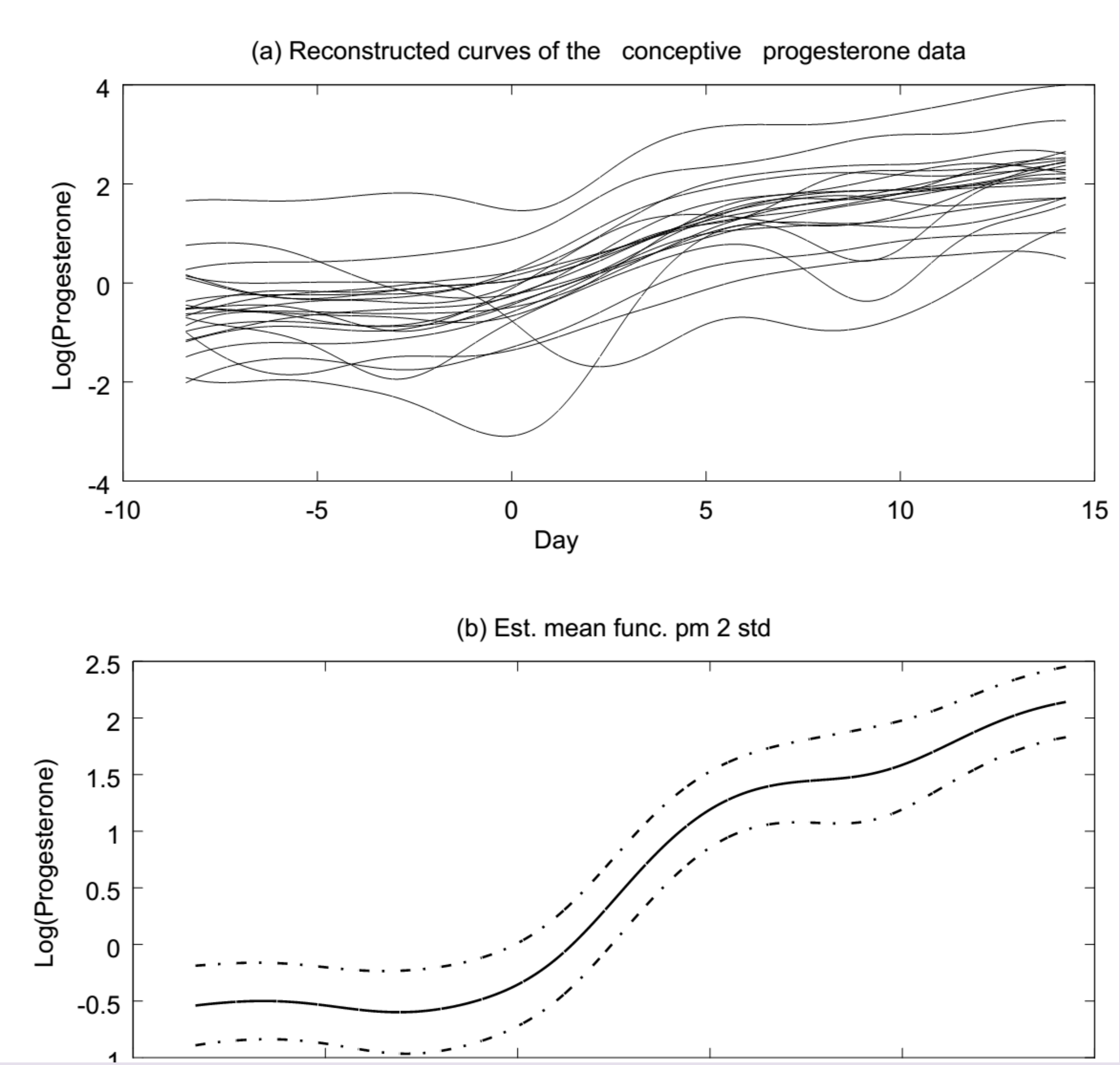
Example 4.14 *Figure 4.11 (a) displays the reconstructed functions of the conceptive progesterone data introduced in Section 1.2.1 of Chapter 1. Their sample mean and covariance functions were computed using the formulas (4.42) given below and are displayed in panels (b) and (c), respectively. From panel (b), it is observed that before the ovulation day (Day 0), the sample mean function of the logarithm of the conceptive progesterone data is near a constant* -0.50. *But after the ovulation day, the sample mean function is no longer a constant; it increases over time. It is then of interest to test the following one-sample problem:*

$$\begin{array}{rll} & H_0 & : \eta(t) \equiv -0.50, \ t \in [a, b], \\ \text{versus} & H_1 & : \eta(t) \neq -0.50, \ \textit{for some } t \in [a, b], \end{array} \tag{4.39}$$

where $[a, b]$ *is any time period of interest. When* $[a, b] = [-8, 0], [0, 15]$ *and* $[-8, 15]$, *we are interested in testing if the underlying mean function of the logarithm of the conceptive progesterone data is a constant* -0.50 *before the ovulation day, after the ovulation day, and over the whole observation period, respectively.*



(a) Reconstructed curves of the nonconceptive progesterone data
Log(Progesterone)
4
2
0
-2
-4
-10
-5
0
5
10
15
Day
(b) Est. mean func. pm 2 std
Log(Progesterone)
2
1.5
1
0.5
0
-0.5
-1
-1.5
-10
-5
0
5
10
15
Day


(a) Reconstructed curves of the conceptive progesterone data
Log(Progesterone)
Day
(b) Est. mean func. pm 2 std
Log(Progesterone)

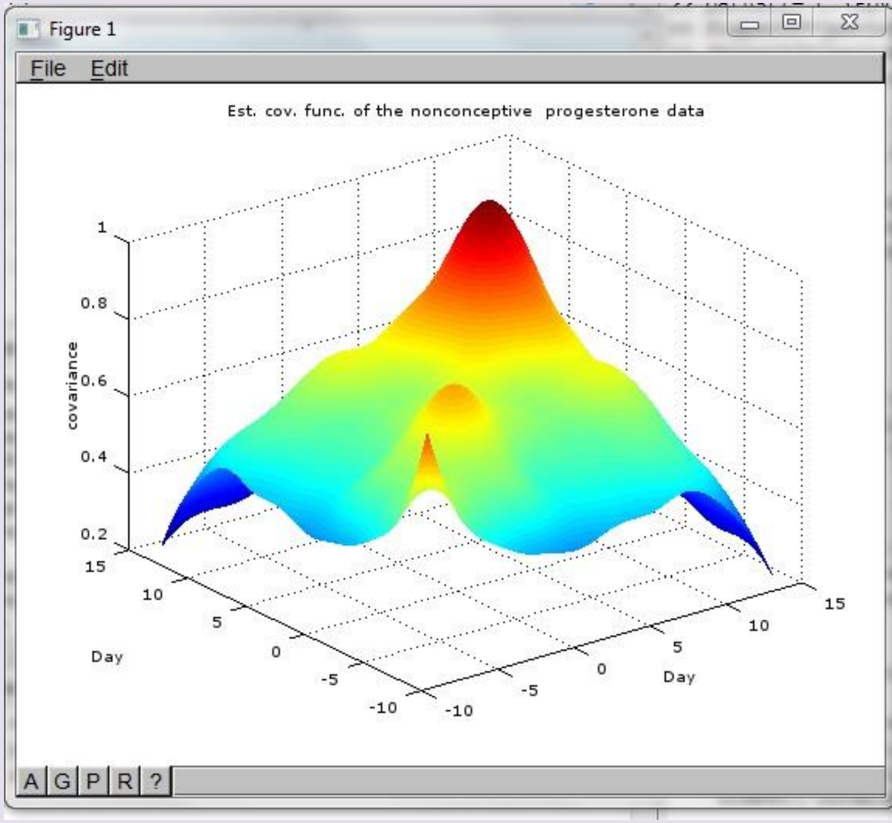


Figure 1
File Edit
Est. cov. func. of the nonconceptive progesterone data
covariance
Day
Day
A G P R ?

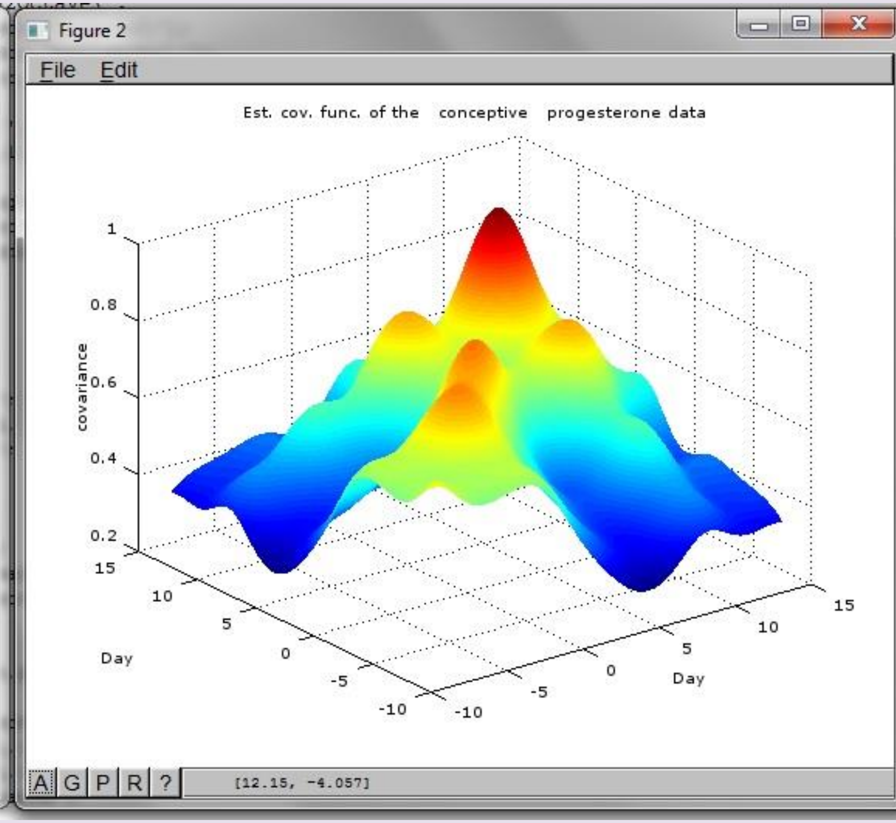


Figure 2
File Edit
Est. cov. func. of the conceptive progesterone data
covariance
Day
Day
A G P R ?
[12.15, -4.057]

GENERAL ONE-SAMPLE PROBLEM

A general one-sample problem for functional data can be described as follows. Suppose we have a functional sample

$$y_1(t), \cdots, y_n(t) \overset{i.i.d.}{\sim} \mathrm{SP}(\eta, \gamma), \tag{4.40}$$

and we wish to test the following hypothesis testing problem:

$$H_0 : \eta(t) \equiv \eta_0(t),\ t \in \mathcal{T}, \quad \text{versus} \quad H_1 : \eta(t) \neq \eta_0(t),\ \text{for some } t \in \mathcal{T}, \tag{4.41}$$

where $\eta_0(t)$ is some known function that is prespecified based on related physical theories, past experiences, or past experimental results. In many situations, $\eta_0(t)$ is specified as 0 to test if the sample is purely noise or if there is some time-effect over $\mathcal{T}$.

UNBIASED ESTIMATOR OF THE MEAN AND COVARIANCE FUNCTIONS

Based on the sample (4.40), the unbiased estimators of $\eta(t)$ and $\gamma(s,t)$, that is, the sample mean and covariance functions, are respectively

$$\begin{aligned} \hat{\eta}(t) &= \bar{y}(t) = n^{-1} \sum_{i=1}^{n} y_i(t), \\ \hat{\gamma}(s,t) &= (n-1)^{-1} \sum_{i=1}^{n} [y_i(s) - \hat{\eta}(s)][y_i(t) - \hat{\eta}(t)]. \end{aligned} \tag{4.42}$$

$$y_1(t), \cdots, y_n(t) \overset{i.i.d.}{\sim} \mathrm{SP}(\eta, \gamma),$$

ONE-SAMPLE PROBLEM ASSUMPTIONS (OS)

1. The functional sample (4.40) is with $\eta(t) \in \mathcal{L}^2(\mathcal{T})$ and $\operatorname{tr}(\gamma) < \infty$.
2. The functional sample (4.40) is Gaussian.
3. The subject-effect function $v_1(t)$ satisfies $\mathrm{E}\|v_1\|^4 = \mathrm{E}\left[\int_{\mathcal{T}} v_1^2(t)dt\right]^2 < \infty$.
4. The maximum variance $\rho = \max_{t \in \mathcal{T}} \gamma(t,t) < \infty$.
5. The expectation $\mathrm{E}[v_1^2(s)v_1^2(t)]$ is uniformly bounded. That is, for any $(s,t) \in \mathcal{T}^2$, we have $\mathrm{E}[v_1^2(s)v_1^2(t)] < C < \infty$, where C is some constant independent of any $(s,t) \in \mathcal{T}^2$.

INSERTION 1

Throughout this book, let $\mathcal{T}$ be a finite interval and we use $\|f\|$ to denote the L^2-norm of a function $f(t), t \in \mathcal{T}$:

$$\|f\| = \left[\int_{\mathcal{T}} f^2(t)dt\right]^{1/2}.$$

If $\|f\| < \infty$, we say that $f(t), t \in \mathcal{T}$ is a squared integrable function. In this case, we write $f(t) \in \mathcal{L}^2(\mathcal{T})$, where $\mathcal{L}^2(\mathcal{T})$ denotes the Hilbert space formed by all the squared integrable functions over $\mathcal{T}$ and the associated inner-product is defined as

$$< f, g >= \int_{\mathcal{T}} f(t)g(t)dt, \quad f(t), g(t) \in \mathcal{L}^2(\mathcal{T}).$$

INSERTION 2

Let $y(t), t \in \mathcal{T}$ be a stochastic process having mean function $\eta(t), t \in \mathcal{T}$ and covariance function $\gamma(s,t), s, t \in \mathcal{T}$, where $\mathcal{T}$ is a compact support of t. We write $y(t) \sim \mathrm{SP}(\eta, \gamma)$ for simplicity. When $\gamma(s,t)$ has finite trace: $\mathrm{tr}(\gamma) = \int_{\mathcal{T}} \gamma(t,t)dt < \infty$, it has the following singular value decomposition (SVD) (Wahba 1990, p. 3):

$$\gamma(s,t) = \sum_{r=1}^{m} \lambda_r \phi_r(s)\phi_r(t), \tag{4.1}$$

$$y_i(t) = \eta(t) + v_i(t) + \epsilon_i(t), \quad i = 1, 2, \cdots, n,$$

INSERTION 3

A process $y(t), t \in \mathcal{T}$ is Gaussian with mean function $\eta(t), t \in \mathcal{T}$ and covariance function $\gamma(s,t), s,t \in \mathcal{T}$, denoted as $\mathrm{GP}(\eta, \gamma)$, if and only if for any p time points, $t_j, j = 1, 2, \cdots, p$, the random vector $[y(t_1), \cdots, y(t_p)]^T$ follows a multivariate normal distribution $N_p(\boldsymbol{\eta}, \boldsymbol{\Gamma})$, where $\boldsymbol{\eta} = [\eta(t_1), \cdots, \eta(t_p)]^T$ and $\boldsymbol{\Gamma} = (\gamma(t_i, t_j)) : p \times p$. Throughout this book, $X \stackrel{d}{=} Y$ denotes that X and Y have the same distribution, and $f(t) \equiv 0, t \in \mathcal{T}$ denotes $f(t) = 0$ for all $t \in \mathcal{T}$.

THEOREM 4.14

1. The functional sample (4.40) is with $\eta(t) \in \mathcal{L}^2(\mathcal{T})$ and $\mathrm{tr}(\gamma) < \infty$.
2. The functional sample (4.40) is Gaussian.

Theorem 4.14 *Under Assumptions OS1 and OS2, we have*

$$\sqrt{n}\,\{\hat{\eta}(t) - \eta(t)\} \sim GP(0, \gamma), \quad (n-1)\hat{\gamma}(s,t) \sim WP(n-1, \gamma).$$

WISHART PROCESS

Wishart processes are natural generalizations of Wishart random matrices. Throughout, we use $\mathrm{WP}(n,\gamma)$ to denote a Wishart process with n degrees of freedom and a covariance function $\gamma(s,t)$. A general Wishart process $W(s,t) \sim \mathrm{WP}(n,\gamma)$ can be written as

$$W(s,t) = \sum_{i=1}^{n} W_i(s,t) = \sum_{i=1}^{n} v_i(s)v_i(t), \tag{4.11}$$

where $W_i(s,t) = v_i(s)v_i(t), i = 1,2,\cdots,n, \overset{i.i.d.}{\sim} \mathrm{WP}(1,\gamma)$ and $v_i(t), i = 1,2,\cdots,n, \overset{i.i.d.}{\sim} \mathrm{GP}(0,\gamma)$. By the definition, we have the following obvious but useful result.

Theorem 4.4 *Let $W_i(s,t) \sim WP(n_i,\gamma)$, $i = 1,2,\cdots,k$. Then we have*

$$W_1(s,t) + W_2(s,t) + \cdots W_k(s,t) \sim WP(n_1 + n_2 + \cdots + n_k, \gamma).$$

PROOF OF THEOREM 4.14

Proof of Theorem 4.14 Under the given conditions, the functional sample (4.40) is Gaussian. The first assertion follows from the fact that $\mathrm{E}\hat{\eta}(t) = \eta(t)$ and $\mathrm{Cov}(\hat{\eta}(s), \hat{\eta}(t)) = \gamma(s,t)/n$. To show the second assertion, let $\mathbf{y}(t) = [y_1(t), y_2(t), \cdots, y_n(t)]^T$. Then $\mathbf{y}(t) \sim \mathrm{GP}_n(\eta \mathbf{1}_n, \gamma \mathbf{I}_n)$. We have

$$(n-1)\hat{\gamma}(s,t) = \mathbf{y}(s)^T(\mathbf{I}_n - \mathbf{J}_n/n)\mathbf{y}(t) = \mathbf{v}(s)^T(\mathbf{I}_n - \mathbf{J}_n/n)\mathbf{v}(t),$$

where $\mathbf{J}_n = \mathbf{1}_n\mathbf{1}_n^T$ is an $n \times n$ matrix of ones and $\mathbf{v}(t) = \mathbf{y}(t) - \eta(t)\mathbf{1}_n \sim \mathrm{GP}_n(\mathbf{0}, \gamma\mathbf{I}_n)$. Notice that $\mathbf{I}_n - \mathbf{J}_n/n$ is an idempotent matrix of rank $n-1$. The second assertion then follows from Theorem 4.8 immediately.

THEOREM 4.8

Theorem 4.8 *Assume* $\mathbf{y}(t) \sim GP_n(\boldsymbol{\eta}, \gamma\mathbf{I}_n)$ *and* $\mathbf{A} : n \times n$ *is an idempotent matrix of rank* k*, having the SVD (4.17). Then the quadratic form* $q(s,t) = \mathbf{y}(s)^T\mathbf{A}\mathbf{y}(t)$ *has the following random expression:*

$$q(s,t) \stackrel{d}{=} \sum_{i=1}^{k} z_i(s)z_i(t), \tag{4.18}$$

where $z_1(t), \cdots, z_n(t)$ *are the entries of* $\mathbf{z}(t) = \mathbf{U}^T\mathbf{y}(t) \sim GP_n(\mathbf{U}^T\boldsymbol{\eta}, \gamma\mathbf{I}_n)$*. In particular, when* $\boldsymbol{\eta}(t) \equiv 0$,

$$q(s,t) \sim WP(k, \gamma). \tag{4.19}$$

THEOREM 4.15

1. The functional sample (4.40) is with $\eta(t) \in \mathcal{L}^2(\mathcal{T})$ and $\mathrm{tr}(\gamma) < \infty$.

Theorem 4.15 *Under Assumption OS1, as $n \to \infty$, we have*

$$\sqrt{n}\,\{\hat{\eta}(t) - \eta(t)\} \xrightarrow{d} GP(0, \gamma),$$

where and throughout, "$\xrightarrow{d}$" denotes "convergence in distribution."

PROOF OF THEOREM 4.15

Proof of Theorem 4.15 Under the given conditions, we have $\mathrm{E}\|y_1\|^2 = \|\eta\|^2 + \mathrm{tr}(\gamma) < \infty$ and the functional sample (4.40) is i.i.d. The assertion follows from Theorem 4.12 immediately.

Theorem 4.12 *If* $y_1(\mathbf{t}), y_2(\mathbf{t}), \cdots, y_n(\mathbf{t}) \stackrel{i.i.d.}{\sim} SP(\eta, \gamma)$ *such that* $E\|y_1\|^2 < \infty$ *where* $\mathbf{t} \in \mathcal{T}$, *then as* $n \to \infty$, *we have*

$$\sqrt{n}[\bar{y}(\mathbf{t}) - \eta(\mathbf{t})] \stackrel{d}{\to} GP(0, \gamma),$$

where $\bar{y}(\mathbf{t}) = n^{-1} \sum_{i=1}^{n} y_i(\mathbf{t})$ *is the usual sample mean function of* $y_i(\mathbf{t}), i = 1, 2, \cdots, n$.

THEOREM 4.16

1. The functional sample (4.40) is with $\eta(t) \in \mathcal{L}^2(\mathcal{T})$ and $\mathrm{tr}(\gamma) < \infty$.

3. The subject-effect function $v_1(t)$ satisfies $\mathrm{E}\|v_1\|^4 = \mathrm{E}\left[\int_{\mathcal{T}} v_1^2(t)dt\right]^2 < \infty$.
4. The maximum variance $\rho = \max_{t \in \mathcal{T}} \gamma(t,t) < \infty$.

Theorem 4.16 *Under Assumptions OS1, OS3, and OS4, as $n \to \infty$, we have*

$$\sqrt{n}\left\{\hat{\gamma}(s,t) - \gamma(s,t)\right\} \xrightarrow{d} GP(0, \varpi),$$

where

$$\varpi\left\{(s_1,t_1),(s_2,t_2)\right\} = E\left\{v_1(s_1)v_1(t_1)v_1(s_2)v_1(t_2)\right\} - \gamma(s_1,t_1)\gamma(s_2,t_2).$$

PROOF OF THEOREM 4.16

Proof of Theorem 4.16 Let $\mathbf{y}(t) = [y_1(t), \cdots, y_n(t)]^T$. Then $\mathbf{y}(t) \sim \mathrm{SP}_n(\eta \mathbf{1}_n, \gamma \mathbf{I}_n)$. Set $\mathbf{v}(t) = \mathbf{y}(t) - \eta(t)\mathbf{1}_n$. Then $\mathbf{v}(t) \sim \mathrm{SP}_n(\mathbf{0}, \gamma \mathbf{I}_n)$. Set $\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n^T$. Then

$$\begin{aligned}\hat{\gamma}(s,t) &= (n-1)^{-1}\mathbf{y}(s)^T(\mathbf{I}_n - \mathbf{J}_n/n)\mathbf{y}(t) \\ &= (n-1)^{-1}\mathbf{v}(s)^T(\mathbf{I}_n - \mathbf{J}_n/n)\mathbf{v}(t) \\ &= (n-1)^{-1}\sum_{i=1}^n z_i(s,t) - \tfrac{n}{n-1}\bar{v}(s)\bar{v}(t),\end{aligned}$$

where $\bar{v}(t) = n^{-1}\sum_{i=1}^n v_i(t)$ and $z_i(s,t) = v_i(s)v_i(t), i = 1, 2, \cdots, n$ are i.i.d. with $\mathrm{E}(z_1(s,t)) = \gamma(s,t)$ and

$$\begin{aligned}\varpi[(s_1,t_1),(s_2,t_2)] &= \mathrm{cov}(z_1(s_1,t_1), z_1(s_2,t_2)) \\ &= \mathrm{E}\left[v_1(s_1)v_1(t_1)v_1(s_2)v_1(t_2)\right] - \gamma(s_1,t_1)\gamma(s_2,t_2).\end{aligned}$$

PROOF OF THEOREM 4.16

By the central limit theorem of i.i.d. stochastic processes, Theorem 4.12, as $n \to \infty$, we have

$$\sqrt{n}\left[n^{-1}\sum_{i=1}^{n} z_i(s,t) - \gamma(s,t)\right] \xrightarrow{d} \mathrm{GP}(0, \varpi),$$

as by Assumption OS3, we have

$$\mathrm{E}\|z_1\|^2 = \mathrm{E}\int_{\mathcal{T}^2}[v_1(s)v_1(t)]^2 ds dt = \mathrm{E}\left[\int_{\mathcal{T}} v_1^2(t)dt\right]^2 = \mathrm{E}\|v_1\|^4 < \infty.$$

It remains to show that $\bar{v}(t) = o_{UP}(1)$, that is, $\bar{v}(t)$ converges to 0 in probability uniformly over $\mathcal{T}$. This is actually true as $\mathrm{E}\bar{v}(t) \equiv 0$ and $\mathrm{Cov}(\bar{v}(s), \bar{v}(t)) = \gamma(s,t)/n \leq \gamma(t,t)/n \leq \rho/n$. Thus under Assumption OS4, we have $\bar{v}(t) = o_{UP}(1)$. The theorem is then proved.

POINTWISE TESTS

GENERAL IDEA

Pointwise tests aim to test the null hypothesis in the one-sample problem (4.41) at each time point $t \in \mathcal{T}$. For any fixed t, the sub-problem is

$$H_{0t} : \eta(t) = \eta_0(t), \quad \text{versus} \quad H_{1t} : \eta(t) \neq \eta_0(t). \tag{4.43}$$

Based on the estimators (4.42), the pivotal test statistic for the above local hypothesis testing problem is

$$z(t) = \frac{\Delta(t)}{\sqrt{\hat{\gamma}(t,t)}} = \frac{\sqrt{n}[\bar{y}(t) - \eta_0(t)]}{\sqrt{\hat{\gamma}(t,t)}}. \tag{4.44}$$

TESTING

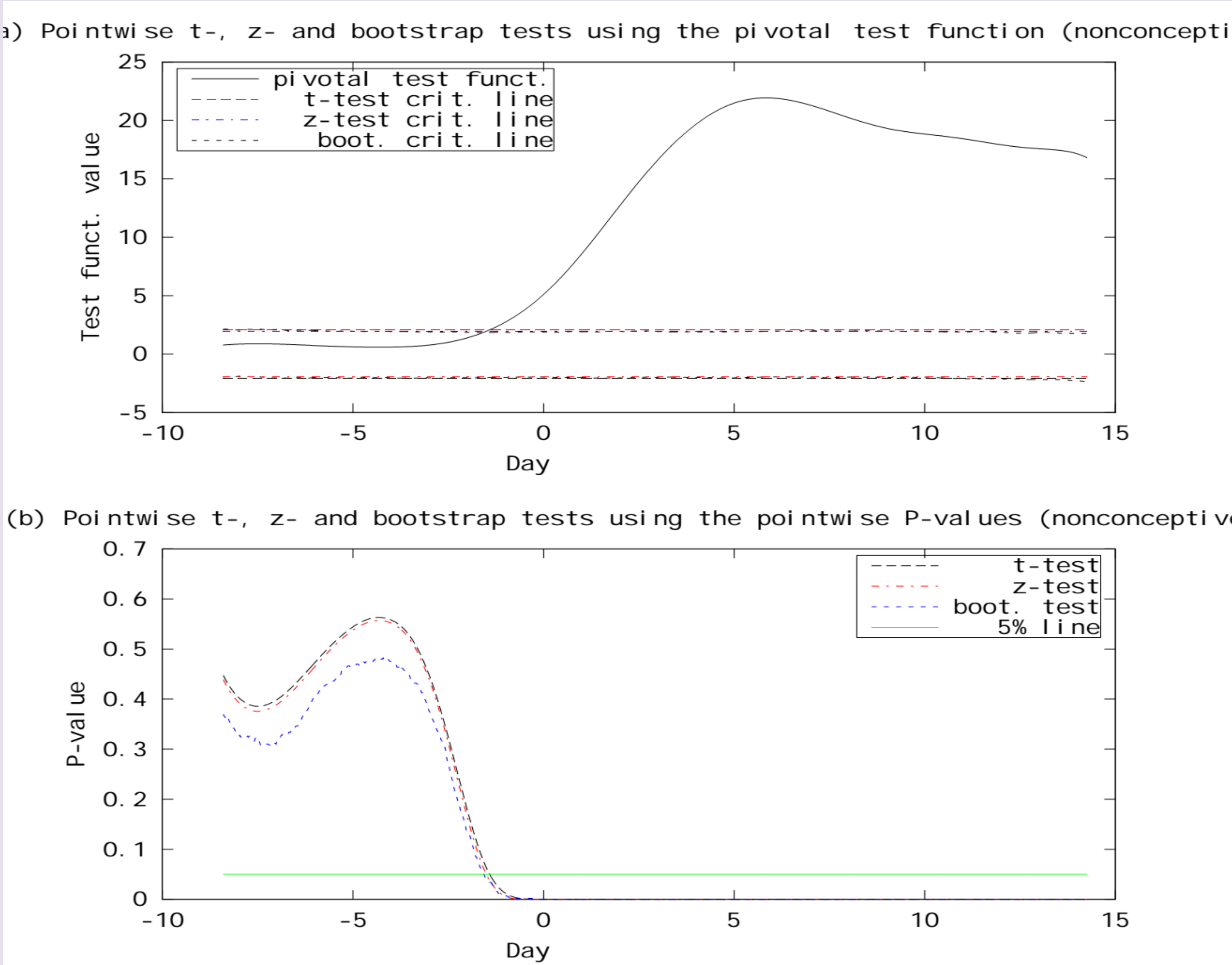
When the functional sample (4.40) is Gaussian, by Theorem 4.14, for each $t \in \mathcal{T}$, we have

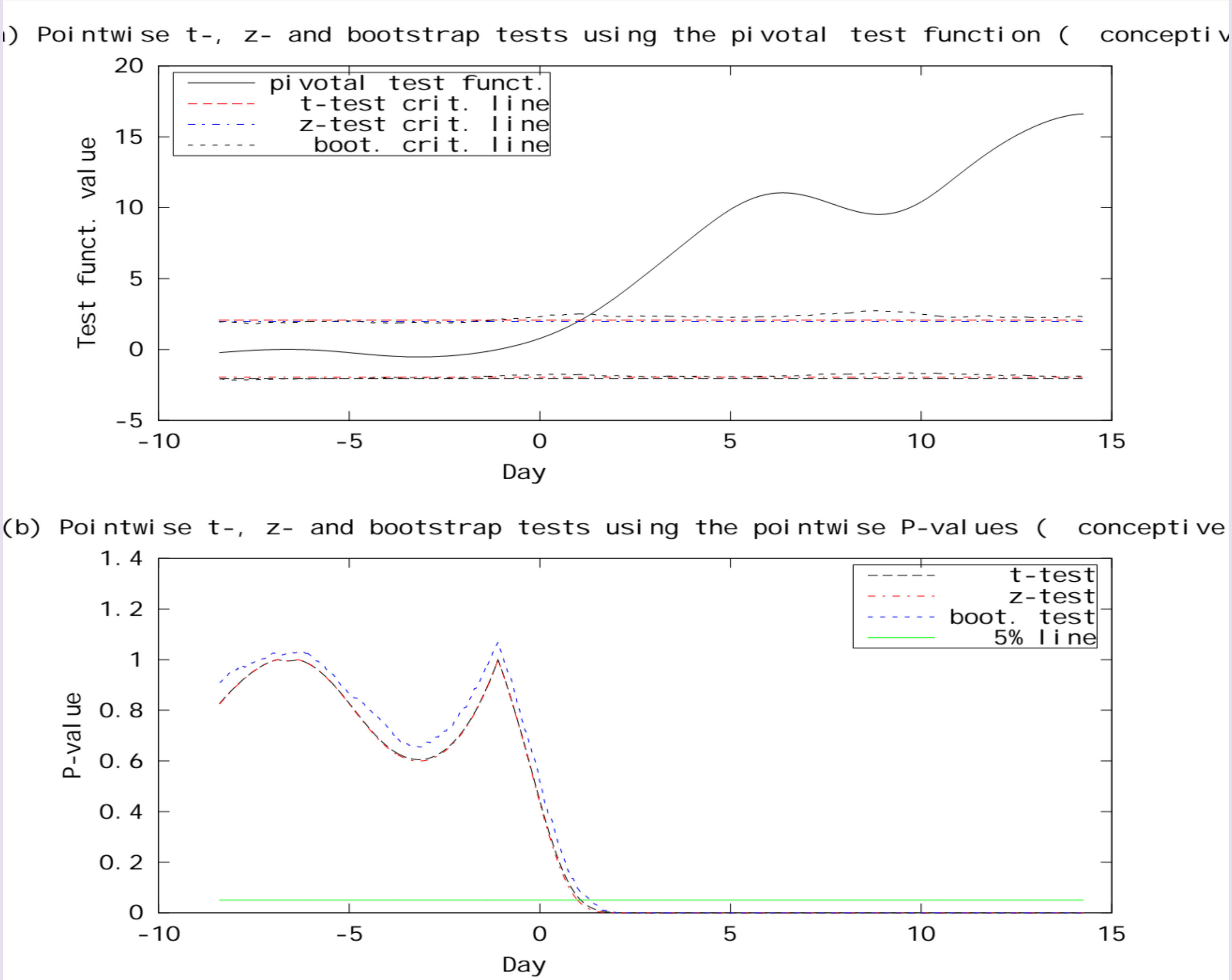
$$z(t) \sim t_{n-1}. \tag{4.45}$$

When the Gaussian assumption is not valid, for large samples, one can use the pointwise z-test. Notice that for any fixed

$t \in \mathcal{T}$, as $n \to \infty$, by Theorem 4.15, we have

$$z(t) \xrightarrow{d} N(0,1), t \in \mathcal{T}. \tag{4.47}$$


a) Pointwise t-, z- and bootstrap tests using the pivotal test function (nonconcepti
pivotal test funct.
t-test crit. line
z-test crit. line
boot. crit. line
Test funct. value
Day
(b) Pointwise t-, z- and bootstrap tests using the pointwise P-values (nonconceptiv
t-test
z-test
boot. test
5% line
P-value
Day


) Pointwise t-, z- and bootstrap tests using the pivotal test function (conceptiv
pivotal test funct.
t-test crit. line
z-test crit. line
boot. crit. line
Test funct. value
Day
20
15
10
5
0
-5
-10
-5
0
5
10
15
(b) Pointwise t-, z- and bootstrap tests using the pointwise P-values (conceptive
t-test
z-test
boot. test
5% line
P-value
Day
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-10
-5
0
5
10
15

L^2 – NORM-BASED TEST

IDEA OF THE TEST

L^2-norm-based test uses the squared L^2-norm of $\Delta(t)$ as its test statistic:

$$T_n = \|\Delta\|^2 = n \int_{\mathcal{T}} [\bar{y}(t) - \eta_0(t)]^2 dt. \qquad (4.50)$$

It is easy to see that T_n will be small under the null hypothesis and it will be large under the alternatives. By Theorem 4.2 and under H_0,

$$T_n \stackrel{d}{=} \sum_{r=1}^{m} \lambda_r A_r, \; A_r \stackrel{i.i.d.}{\sim} \chi_1^2,$$

which is valid when the Gaussian assumption holds or is asymptotically valid when n is large.

The null distribution of T_n can then be approximated using the methods described in Section 4.3. For example, by the Welch-Satterthwaite χ^2-approximation, we have

$$T_n \sim \beta\chi_d^2 \text{ approximately,} \quad \text{where } \beta = \frac{\text{tr}(\gamma^{\otimes 2})}{\text{tr}(\gamma)},\ d = \kappa = \frac{\text{tr}^2(\gamma)}{\text{tr}(\gamma^{\otimes 2})}. \tag{4.51}$$

Remark 4.10 *The parameters β and κ depend on the underlying covariance function $\gamma(s,t)$ only. They not only depend on the variances $\gamma(t,t), t \in \mathcal{T}$ via $tr(\gamma)$, but also depend on the covariances $\gamma(s,t),\ s \neq t,\ s,t \in \mathcal{T}$ by $tr(\gamma^{\otimes 2})$. In this sense, the L^2-norm-based test and the F-type test defined in the next subsection do partially take into account the dependence of functional data. The parameter β is proportional to the scale of the functional data while the parameter κ is scale invariant. Later in this book we shall see that the approximate degrees of freedom of many test statistics for functional data are proportional to κ. That is why Shen and Faraway (2004) called κ the degrees of freedom adjustment factor.*

DEFINITIONS

$$
\begin{aligned}
(f \otimes g)(s,t) &= \int_{\mathcal{T}} f(s,u) g(u,t) du, \\
f^{\otimes 2}(s,t) &= \int_{\mathcal{T}} f(s,u) f(u,t) du.
\end{aligned} \tag{4.5}
$$

$$
\operatorname{tr}(\gamma) = \int_{\mathcal{T}} \gamma(t,t) dt < \infty,
$$

NAIVE ESTIMATOR

In practice, the parameters β and κ must be estimated based on the functional data (4.40). A natural way to do this is to replace $\gamma(s,t)$ in $\text{tr}(\gamma), \text{tr}^2(\gamma)$ and $tr(\gamma^{\otimes 2})$ with its unbiased estimator $\hat{\gamma}(s,t)$ so that

$$\hat{\beta} = \frac{\text{tr}(\hat{\gamma}^{\otimes 2})}{\text{tr}(\hat{\gamma})}, \quad \hat{d} = \hat{\kappa} = \frac{\text{tr}^2(\hat{\gamma})}{\text{tr}(\hat{\gamma}^{\otimes 2})}, \tag{4.52}$$

where $\hat{\gamma}(s,t)$ is given in (4.42) based on the sample (4.40). In this case, we have

$$T_n \sim \hat{\beta}\chi^2_{\hat{d}} \text{ approximately.} \tag{4.53}$$

BIAS-REDUCED ESTIMATOR

$$\widehat{\mathrm{tr}^2(\gamma)} = \frac{(n-1)n}{(n-2)(n+1)}\left[\mathrm{tr}^2(\hat{\gamma}) - \frac{2\mathrm{tr}(\hat{\gamma}^{\otimes 2})}{n}\right],$$

$$\mathrm{tr}\widehat{(\gamma^{\otimes 2})} = \frac{(n-1)^2}{(n-2)(n+1)}\left[\mathrm{tr}(\hat{\gamma}^{\otimes 2}) - \frac{\mathrm{tr}^2(\hat{\gamma})}{n-1}\right]. \quad (4.54)$$

Replacing $\mathrm{tr}^2(\gamma)$ and $\mathrm{tr}(\gamma^{\otimes 2})$ in (4.51) by their unbiased estimator defined above results in the so-called bias-reduced method for estimating the parameters β and d:

$$\hat{\beta} = \frac{\mathrm{tr}\widehat{(\gamma^{\otimes 2})}}{\mathrm{tr}(\hat{\gamma})}, \quad \hat{d} = \hat{\kappa} = \frac{\widehat{\mathrm{tr}^2(\gamma)}}{\mathrm{tr}\widehat{(\gamma^{\otimes 2})}}. \quad (4.55)$$

PROPERTIES OF THE ESTIMATORS

1. The functional sample (4.40) is with $\eta(t) \in \mathcal{L}^2(\mathcal{T})$ and $\mathrm{tr}(\gamma) < \infty$.

3. The subject-effect function $v_1(t)$ satisfies $\mathrm{E}\|v_1\|^4 = \mathrm{E}\left[\int_{\mathcal{T}} v_1^2(t)dt\right]^2 < \infty$.
4. The maximum variance $\rho = \max_{t\in\mathcal{T}} \gamma(t,t) < \infty$.
5. The expectation $\mathrm{E}[v_1^2(s)v_1^2(t)]$ is uniformly bounded. That is, for any $(s,t) \in \mathcal{T}^2$, we have $\mathrm{E}[v_1^2(s)v_1^2(t)] < C < \infty$, where C is some constant independent of any $(s,t) \in \mathcal{T}^2$.

Theorem 4.17 *Under Assumptions OS1 and OS3 through OS5, as $n \to \infty$, we have $tr(\hat{\gamma}) \xrightarrow{p} tr(\gamma)$ and $tr(\hat{\gamma}^{\otimes 2}) \xrightarrow{p} tr(\gamma^{\otimes 2})$. Furthermore, as $n \to \infty$, we have $\hat{\beta} \xrightarrow{p} \beta$ and $\hat{\kappa} \xrightarrow{p} \kappa$, where $\hat{\beta}$ and $\hat{\kappa}$ are the naive or bias-reduced estimators of β and κ.*

Proof of Theorem 4.17 Under the given conditions, by Theorem 4.16, as $n \to \infty$, we have $\mathrm{E}[\hat{\gamma}(s,t) - \gamma(s,t)]^2 = \frac{\omega[(s,t),(s,t)]}{n}[1 + o(1)]$. By Assumptions OS4 and OS5, we have

$$|\omega[(s,t),(s,t)]| \le \mathrm{E}[v_1^2(s)v_1^2(t)] + \gamma^2(s,t) \le C + \rho, \quad \text{for all } (s,t) \in \mathcal{T}^2.$$

Thus, we have $\hat{\gamma}(s,t) = \gamma(s,t) + O_{UP}(n^{-1/2}),\ (s,t) \in \mathcal{T}^2$, where O_{UP} means "uniformly bounded in probability." It follows that $\hat{\gamma}(s,t) \xrightarrow{p} \gamma(s,t)$ uniformly over $\mathcal{T}^2$. Therefore,

$$\begin{aligned}
\lim_{n\to\infty} \mathrm{tr}(\hat{\gamma}) &= \int_{\mathcal{T}} \lim_{n\to\infty} \hat{\gamma}(t,t)dt = \int_{\mathcal{T}} \gamma(t,t)dt = \mathrm{tr}(\gamma), \\
\lim_{n\to\infty} \mathrm{tr}(\hat{\gamma}^{\otimes 2}) &= \int_{\mathcal{T}} \int_{\mathcal{T}} \lim_{n\to\infty} \hat{\gamma}^2(s,t)dsdt \\
&= \int_{\mathcal{T}} \int_{\mathcal{T}} \gamma^2(s,t)dsdt = \mathrm{tr}(\gamma^{\otimes 2}).
\end{aligned}$$

It follows from (4.52) and (4.54) that as $n \to \infty$, $\hat{\beta} \xrightarrow{p} \beta$ and $\hat{\kappa} \xrightarrow{p} \kappa$. The theorem is proved.

EXAMPLE

	Method	[a,b]	T_n	$\widehat{\beta}$	$\widehat{d} = \hat{\kappa}$	P-value
Conceptive	Naive	[-8,0]	28.48	235.26	1.13	0.773
		[0,15]	46967	370.74	1.24	0
		[-8,15]	46996	460.81	1.58	0
	Bias-reduced	[-8,0]	28.48	213.45	1.15	0.768
		[0,15]	46967	334.38	1.28	0
		[-8,15]	46996	408.63	1.68	0
Nonconceptive	Naive	[-8,0]	752.54	202.99	1.24	0.075
		[0,15]	15962	381.03	1.25	0
		[-8,15]	16037	368.75	1.44	0
	Bias-reduced	[-8,0]	752.54	196.48	1.25	0.071
		[0,15]	15962	505.70	1.26	0
		[-8,15]	16037	488.02	1.46	0

F-TYPE TEST

IDEA OF THE TEST

Then it is natural to test (4.41) using the following test statistic:

$$F_n = \frac{\|\Delta\|^2}{\mathrm{tr}(\hat{\gamma})} = \frac{n \int_{\mathcal{T}} [\bar{y}(t) - \eta_0(t)]^2 dt}{\mathrm{tr}(\hat{\gamma})}. \quad (4.57)$$

$H_0 : \eta(t) \equiv \eta_0(t),\ t \in \mathcal{T}, \quad$ versus $\quad H_1 : \eta(t) \neq \eta_0(t),$ for some $t \in \mathcal{T}$,

IDEA OF THE TEST

When the variation of $\text{tr}(\hat{\gamma})$ is not taken into account, the above test statistic F_n is equivalent to the L^2-norm-based test statistic T_n defined in (4.50).

$$F_n \sim F_{1,n-1}$$

$$F_n \sim F_{\hat{\kappa},(n-1)\hat{\kappa}} \text{ approximately,}$$

$$\hat{d} = \hat{\kappa} = \frac{\text{tr}^2(\hat{\gamma})}{\text{tr}(\hat{\gamma}^{\otimes 2})},$$

$$\hat{d} = \hat{\kappa} = \frac{\widehat{\text{tr}^2(\gamma)}}{\text{tr}\widehat{(\gamma^{\otimes 2})}}.$$

$$\hat{\gamma}(s,t) \quad = \quad (n-1)^{-1} \sum_{i=1}^{n} [y_i(s) - \hat{\eta}(s)][y_i(t) - \hat{\eta}(t)].$$

EXAMPLE

	Method	[a,b]	F test	$\hat{\kappa}$	$(n-1)\hat{\kappa}$	P-value
Conceptive	Naive	[-8,0]	0.107	1.13	23.65	0.776
		[0,15]	101.85	1.24	26.12	2.6578e-011
		[-8,15]	64.73	1.58	33.09	3.8508e-011
	Bias-reduced	[-8,0]	0.107	1.16	24.06	0.776
		[0,15]	101.85	1.28	26.96	1.3507e-011
		[-8,15]	64.73	1.68	35.31	9.5934e-012
Nonconceptive	Naive	[-8,0]	2.99	1.24	84.29	0.079
		[0,15]	334.74	1.25	85.10	0
		[-8,15]	220.15	1.44	97.95	0
	Bias-reduced	[-8,0]	2.99	1.25	85.08	0.078
		[0,15]	334.74	1.26	85.93	0
		[-8,15]	220.15	1.46	99.50	0

BOOTSTRAP TEST

GENERAL IDEA

$$y_1(t), \cdots, y_n(t) \stackrel{i.i.d.}{\sim} \mathrm{SP}(\eta, \gamma),$$

$$y_i^*(t), i = 1, 2, \cdots, n,$$

Based on the bootstrap sample (4.49) randomly generated from (4.40), we can construct the sample mean and sample covariance functions as $\bar{y}^*(t)$ and $\hat{\gamma}^*(s,t)$, computed as in (4.42).

$$\begin{aligned} \hat{\eta}(t) &= \bar{y}(t) = n^{-1} \textstyle\sum_{i=1}^n y_i(t), \\ \hat{\gamma}(s,t) &= (n-1)^{-1} \textstyle\sum_{i=1}^n [y_i(s) - \hat{\eta}(s)][y_i(t) - \hat{\eta}(t)]. \end{aligned}$$

L^2 – NORM – BASED BOOTSTRAP TEST

The L^2-norm-based bootstrap test computes

the bootstrap test statistic $T_n^* = \|\Delta^*\|$ with the bootstrapped pivotal test function $\Delta^*(t) = \sqrt{n}\,[\bar{y}^*(t) - \bar{y}(t)]$. Repeat this process a large number of times so that one can obtain a bootstrap sample of T_n^* that can be used to

estimate the $100(1-\alpha)$-percentile of T_n.

F-TYPE BOOTSTRAP TEST

Similarly, the F-type bootstrap test computes the bootstrap test statistic $F_n^* = \frac{\|\Delta^*\|}{\mathrm{tr}(\hat{\gamma}^*)}$. Repeat this process a large number of times so that one can obtain a bootstrap sample of F_n^* that allows to estimate the $100(1-\alpha)$-percentile of F_n.

EXAMPLE

	Method	[a,b]	L_2 test	P-value	F test	P-value
Conceptive	Naive	[-8,0]	28.48	0.828	0.107	0.833
		[0,15]	46967	0	101.85	0
		[-8,15]	46996	0	64.73	0
Nonconceptive	Naive	[-8,0]	752.54	0.076	2.99	0.085
		[0,15]	15962	0	334.74	0
		[-8,15]	16037	0	220.15	0

NUMERICAL IMPLEMENTATION

DISCRETIZATION

In practice, we have to discretize the continuous functions $\Delta(t)$ and $\hat{\gamma}(s,t)$ in the computation of $\|\Delta\|^2, \mathrm{tr}(\hat{\gamma})$ and $\mathrm{tr}(\hat{\gamma}^{\otimes 2})$. Let the resolution number be M, a large number, say, $M = 1,000$, and let $t_1, t_2, \cdots, t_M$ be M resolution time points that are equally spaced in $\mathcal{T}$. Then the functional sample (4.40) is discretized accordingly as

$$\mathbf{y}_i = [y_i(t_1), y_i(t_2), \cdots, y_i(t_M)]^T, \ i = 1, 2, \cdots, n, \tag{4.58}$$

and the one-sample problem (4.41) is discretized as

$$H_0 : \boldsymbol{\eta} = \boldsymbol{\eta}_0, \quad \text{versus} \quad H_1 : \boldsymbol{\eta} \neq \boldsymbol{\eta}_0, \tag{4.59}$$

where $\boldsymbol{\eta} = [\eta(t_1), \cdots, \eta(t_M)]^T$ and $\boldsymbol{\eta}_0 = [\eta_0(t_1), \cdots, \eta_0(t_M)]^T$

DISCRETIZATION

$$\begin{aligned} T_n &= \|\Delta\|^2 = \int_{\mathcal{T}} \Delta^2(t)dt \approx \frac{v(\mathcal{T})}{M} \sum_{i=1}^M \Delta^2(t_i) \\ &= \frac{v(\mathcal{T})}{M} \|\boldsymbol{\Delta}\|^2 = \frac{v(\mathcal{T})}{M} T_n^0, \end{aligned} \tag{4.61}$$

where $v(\mathcal{T})$ denotes the volume of $\mathcal{T}$ and $T_n^0 = \|\boldsymbol{\Delta}\|^2$ denotes the usual squared L^2-norm of $\boldsymbol{\Delta}$. When $\mathcal{T} = [a, b]$, one has $v(\mathcal{T}) = b - a$. Similarly, we have

$$\operatorname{tr}(\hat{\gamma}) = \int_{\mathcal{T}} \hat{\gamma}(t,t)dt \approx \frac{v(\mathcal{T})}{M} \sum_{i=1}^M \hat{\gamma}(t_i, t_i) = \frac{v(\mathcal{T})}{M} \operatorname{tr}(\hat{\boldsymbol{\Gamma}}). \tag{4.62}$$

In addition, we have

$$\operatorname{tr}(\hat{\gamma}^{\otimes 2}) = \int_{\mathcal{T}^2} \hat{\gamma}^2(s,t)dsdt \approx \frac{v(\mathcal{T}^2)}{M^2} \sum_{i=1}^M \sum_{j=1}^M \hat{\gamma}^2(t_i, t_j) = \frac{v(\mathcal{T}^2)}{M^2} \operatorname{tr}(\hat{\boldsymbol{\Gamma}}^2).$$

When $\mathcal{T} = [a, b]$, which is often the case, one has $v(\mathcal{T}^2) = v^2(\mathcal{T}) = (b-a)^2$.

NOTATIONS

$$\mathbf{\Delta} = \begin{bmatrix} \Delta(t_1) \\ \Delta(t_2) \\ \vdots \\ \Delta(t_M) \end{bmatrix}, \quad \text{and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} \hat{\gamma}(t_1, t_1) & \cdots & \hat{\gamma}(t_1, t_M) \\ \hat{\gamma}(t_2, t_1) & \cdots & \hat{\gamma}(t_2, t_M) \\ \vdots & \vdots & \vdots \\ \hat{\gamma}(t_M, t_1) & \cdots & \hat{\gamma}(t_M, t_M) \end{bmatrix}, \qquad (4.60)$$

DISCRETIZATION

this case, the estimated parameters $\hat{\beta}$ and $\hat{\kappa}$ defined in (4.52) for the Welch-Satterthwaite χ^2-approximation can be approximately expressed as

$$\hat{\beta} \approx \frac{v(\mathcal{T})}{M} \frac{\mathrm{tr}(\hat{\mathbf{\Gamma}}^2)}{\mathrm{tr}(\hat{\mathbf{\Gamma}})} = \frac{v(\mathcal{T})}{M} \hat{\beta}^0, \quad \text{and} \quad \hat{\kappa} \approx \frac{\mathrm{tr}^2(\hat{\mathbf{\Gamma}})}{\mathrm{tr}(\hat{\mathbf{\Gamma}}^2)} = \hat{\kappa}^0,$$

where $\hat{\beta}^0$ and $\hat{\kappa}^0$ denote the estimated β and κ when the Welch-Satterthwaite χ^2-approximation is applied to the discretized one-sample problem (4.59) based on the discretized one sample (4.58). It follows that

$$P\left(T_n \geq \hat{\beta}\chi^2_{\hat{\kappa}}\right) \approx P\left(\frac{v(\mathcal{T})}{M} T_n^0 \geq \frac{v(\mathcal{T})}{M} \hat{\beta}^0 \chi^2_{\hat{\kappa}^0}\right) = P\left(T_n^0 \geq \hat{\beta}^0 \chi^2_{\hat{\kappa}^0}\right).$$

We then have the following remark.

REMARK

Remark 4.13 *When we conduct the L^2-norm-based test, the constant factor $\frac{v(\mathcal{T})}{M}$ in T_n and $\hat{\beta}$ can be omitted at the same time in computation. This will not affect the test result.*

DISCRETIZATION

The above remark is also true for the F-type test. In fact, by (4.61) and (4.62), the F-type test statistic (4.57) can be approximately expressed as

$$F_n = \frac{\|\Delta\|^2}{\mathrm{tr}(\hat{\gamma})} \approx \frac{\|\mathbf{\Delta}\|^2}{\mathrm{tr}(\hat{\mathbf{\Gamma}})} \equiv F_n^0, \tag{4.63}$$

where F_n^0 denotes the test statistic for the F-type test applied to the discretized one-sample problem (4.59) based on the discretized one sample (4.58) so that

$$P(F_n \geq F_{\hat{\kappa},(n-1)\hat{\kappa}}) \approx P(F_n^0 \geq F_{\hat{\kappa}^0,(n-1)\hat{\kappa}^0}).$$

REMARK

Remark 4.14 *From the above, it is seen that in practice, we have to discretize the functional data so that the various tests described in this section are actually applied to the discretized sample (4.58) by computing* $T_n^0, F_n^0, tr(\hat{\mathbf{\Gamma}}), tr(\hat{\mathbf{\Gamma}}^2), \hat{\beta}^0, \hat{\kappa}^0$, *etc. Therefore, when the functional data are very noisy and cannot be reconstructed by the methods proposed in Chapter 3 but they can be observed simultaneously over a common grid of time points so that a sample of vectors like (4.58) can be obtained, then the various tests described in this section can be applied directly to the observed functional data. This is true for all the methodologies investigated in this book.*

REMARK

Remark 4.15 *When M is large and n is relatively small, we can simplify the computation using the following technique. Note that*

$$\hat{\gamma}(s,t) = (n-1)^{-1}\sum_{i=1}^{n}\hat{v}_i(s)\hat{v}_i(t) = (n-1)^{-1}\hat{\mathbf{v}}(s)^T\hat{\mathbf{v}}(t),$$

where $\hat{\mathbf{v}}(t) = [\hat{v}_1(t), \cdots, \hat{v}_n(t)]^T$ with $\hat{v}_i(t) = y_i(t) - \bar{y}(t), i = 1, 2, \cdots, n$ being the estimated subject-effect functions. It follows that

$$\hat{\mathbf{\Gamma}} = (n-1)^{-1}\hat{\mathbf{V}}\hat{\mathbf{V}}^T,$$

where $\hat{\mathbf{V}} = [\hat{\mathbf{v}}(t_1), \hat{\mathbf{v}}(t_2), \cdots, \hat{\mathbf{v}}(t_M)]^T : M \times n$. Thus

$$\begin{aligned} tr(\hat{\mathbf{\Gamma}}) &= (n-1)^{-1}tr\left(\hat{\mathbf{V}}\hat{\mathbf{V}}^T\right) = (n-1)^{-1}tr(\mathbf{S}), \\ tr(\hat{\mathbf{\Gamma}}^2) &= (n-1)^{-2}tr\left[(\hat{\mathbf{V}}\hat{\mathbf{V}}^T)^2\right] = (n-1)^{-2}tr(\mathbf{S}^2), \end{aligned}$$

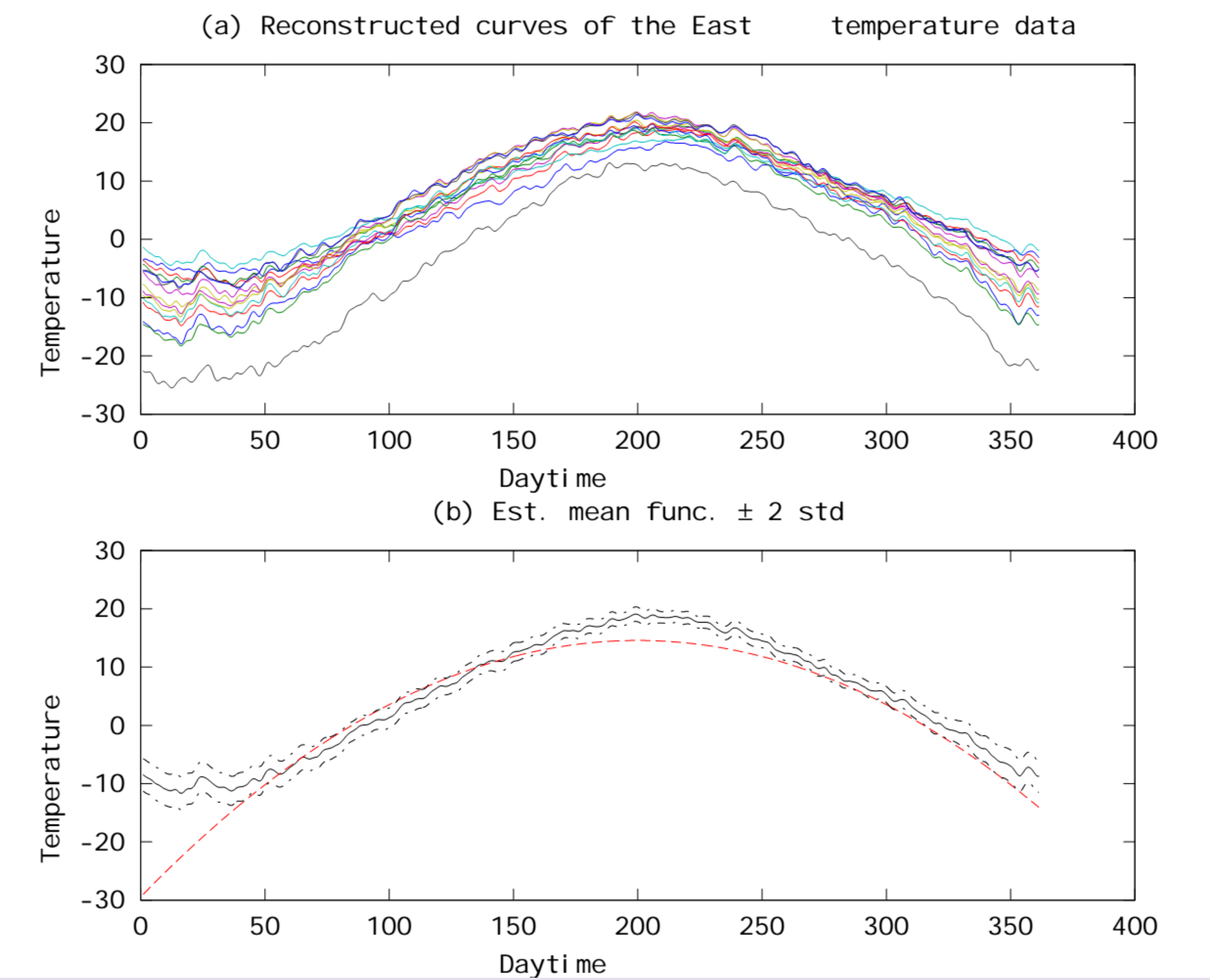
where $\mathbf{S} = \hat{\mathbf{V}}^T\hat{\mathbf{V}}$ is an $n \times n$ matrix so that the needed operations for computing $tr(\hat{\mathbf{\Gamma}})$ and $tr(\hat{\mathbf{\Gamma}}^2)$ are $O(n)$ and $O(n^2)$ instead of $O(M)$ and $O(M^2)$, respectively. This saves a lot of computation.

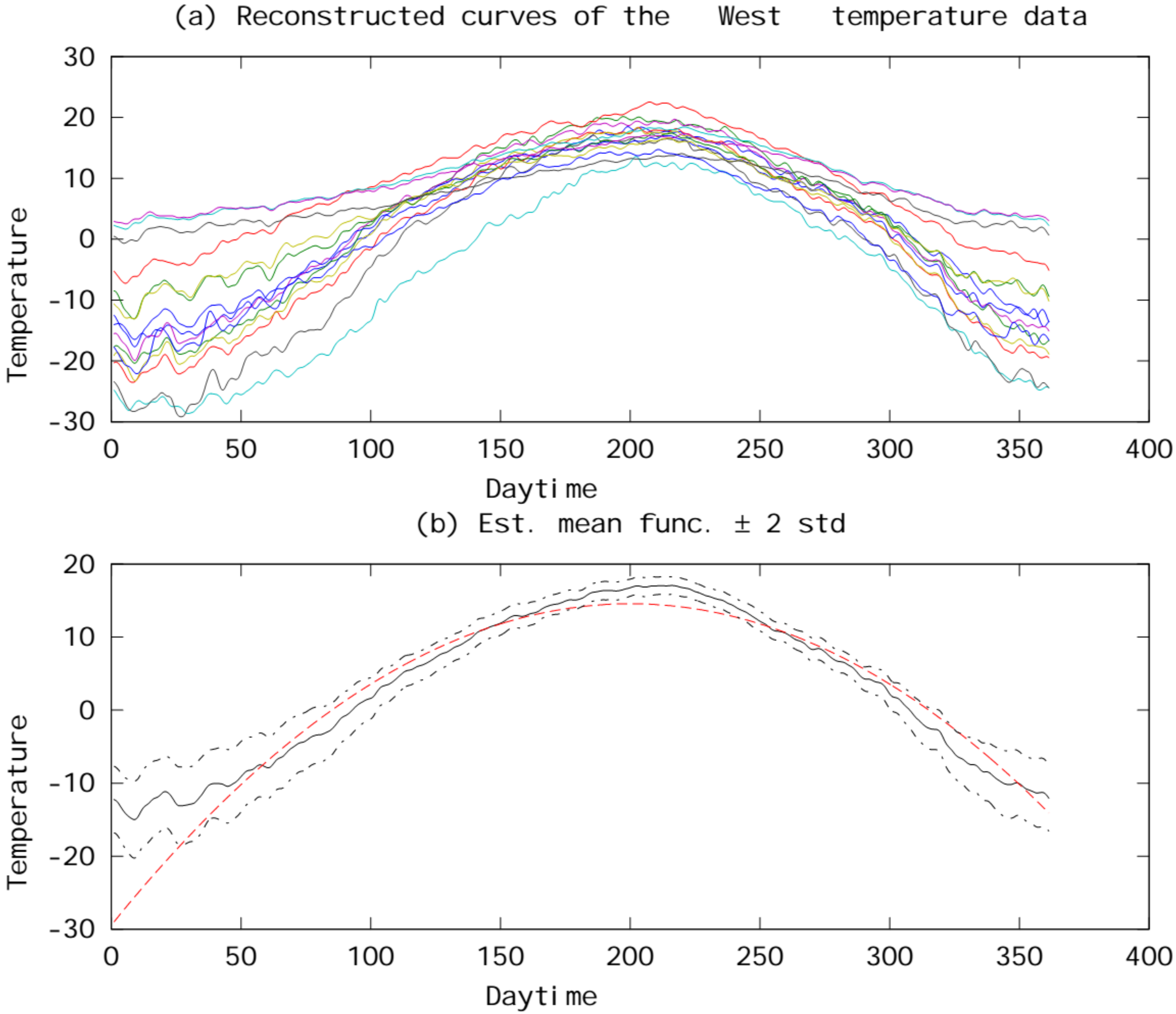
TEMPERATURE EXAMPLE

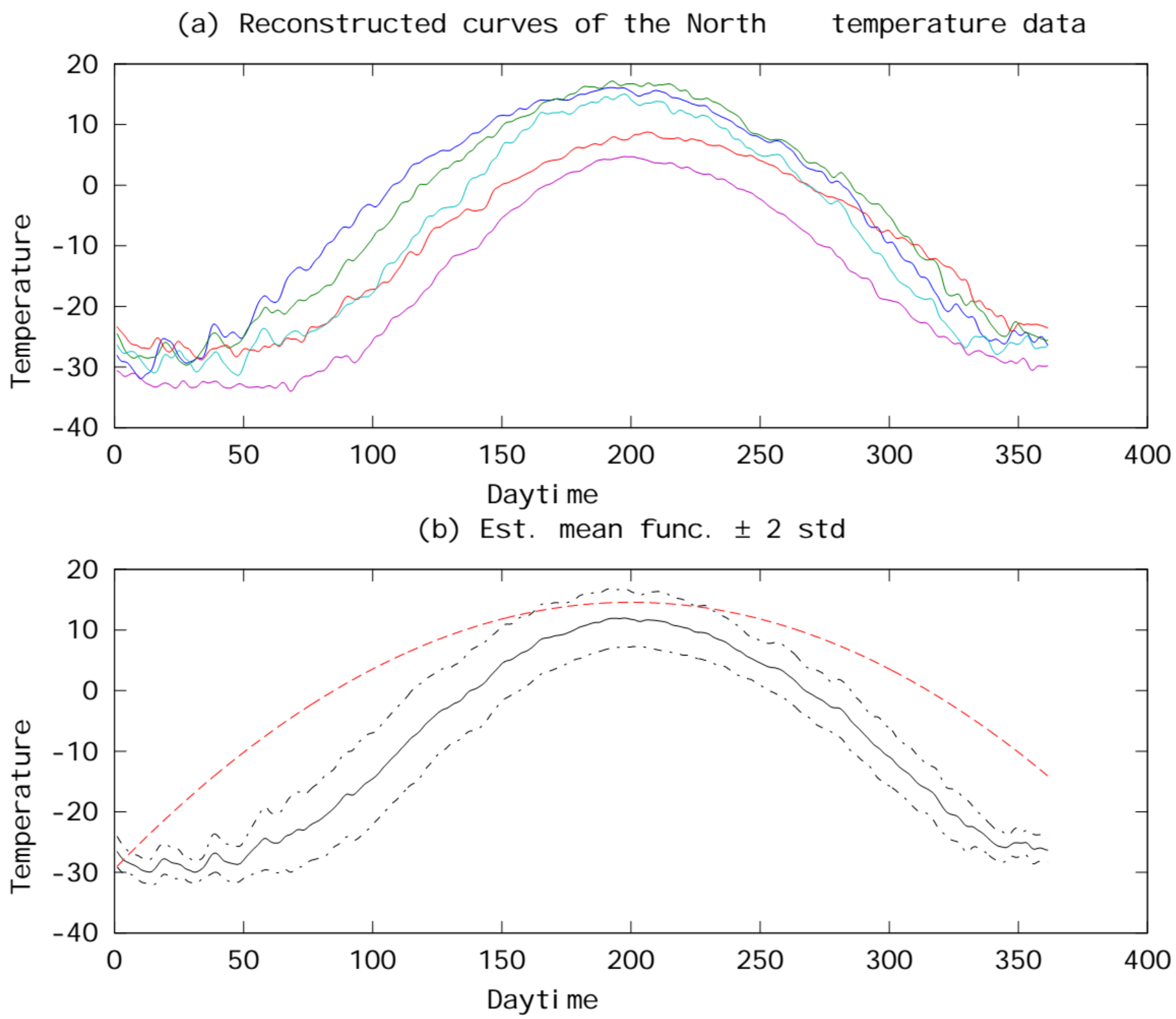
We want to test a hypothesis about the mean of Canadian Temperature Data

$$H_0: \mu = -0.0011t^2 + 0.44t - 29.44,$$
$$t \in [0,365]$$
$$H_1: \mu \neq -0.0011t^2 + 0.44t - 29.44,$$
$$for\ some\ t \in [0,365]$$

We investigate this hypothesis in three different regions: North; East; West.


(a) Reconstructed curves of the East temperature data
Temperature
Daytime
(b) Est. mean func. ± 2 std
Temperature
Daytime


(a) Reconstructed curves of the West temperature data
30
20
10
0
-10
-20
-30
Temperature
0
50
100
150
200
250
300
350
400
Daytime
(b) Est. mean func. ± 2 std
20
10
0
-10
-20
-30
Temperature
0
50
100
150
200
250
300
350
400
Daytime



(a) Reconstructed curves of the North temperature data
Temperature
Daytime
(b) Est. mean func. ± 2 std
Temperature
Daytime

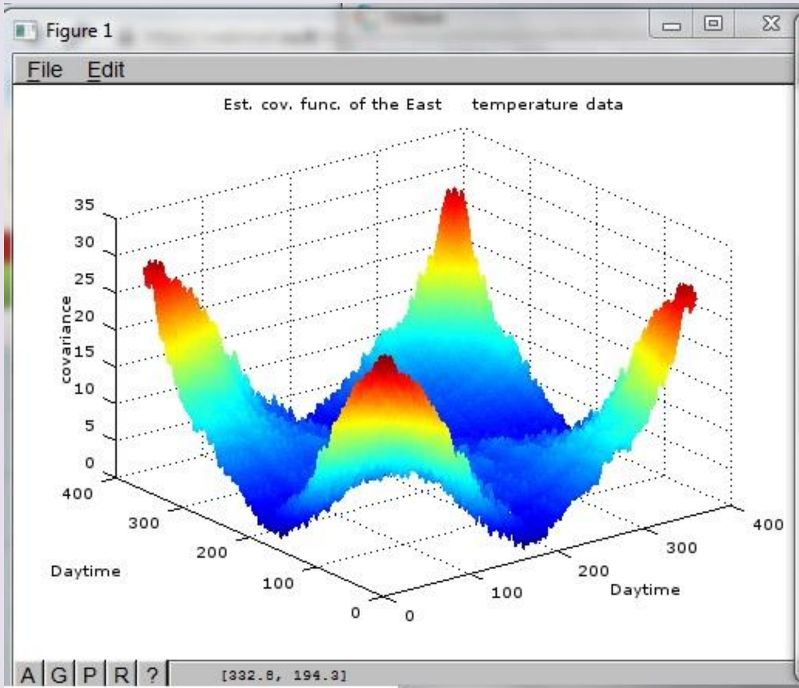


Figure 1
File Edit
Est. cov. func. of the East temperature data
covariance
Daytime
Daytime
A G P R ?
[332.8, 194.3]

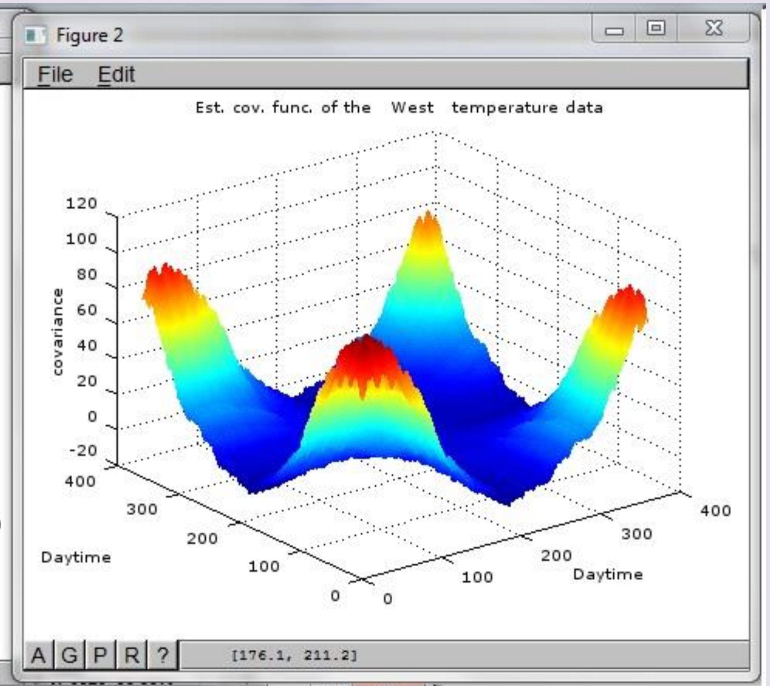


Figure 2
File Edit
Est. cov. func. of the West temperature data
covariance
Daytime
Daytime
A G P R ?
[176.1, 211.2]

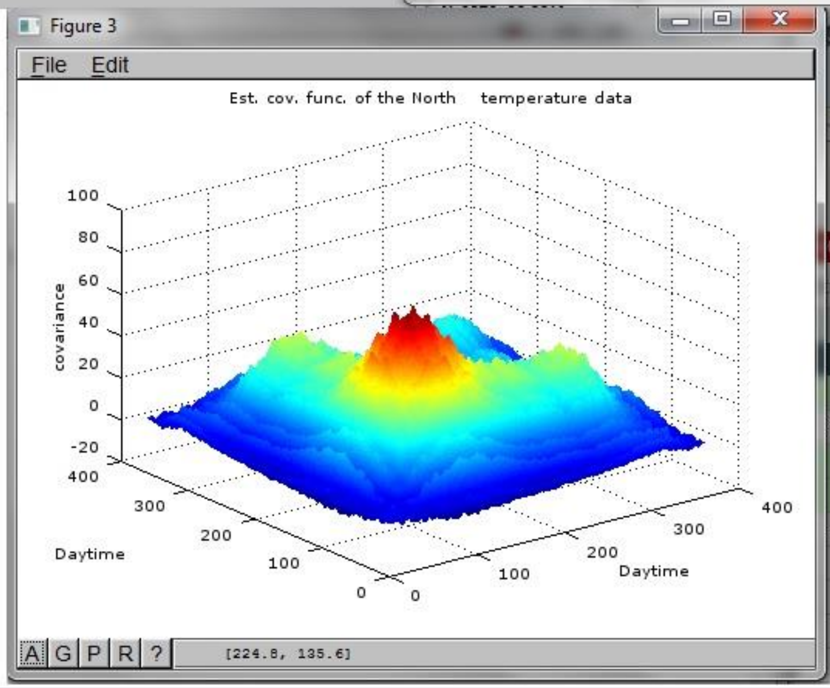
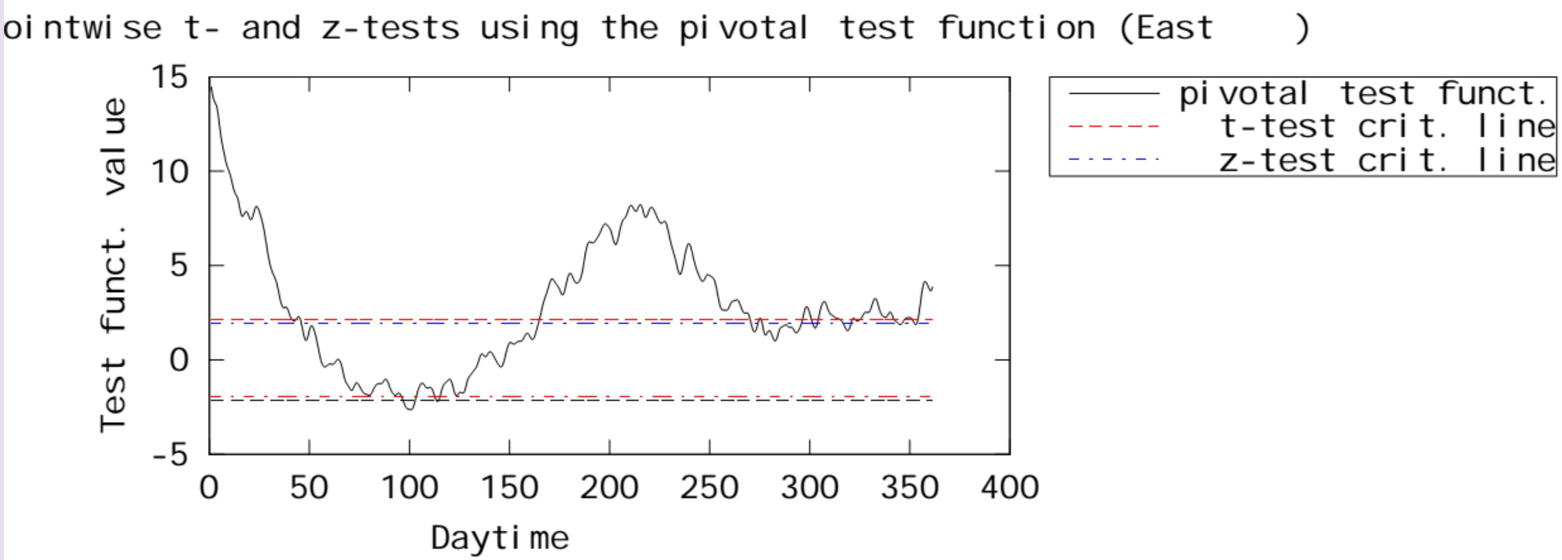
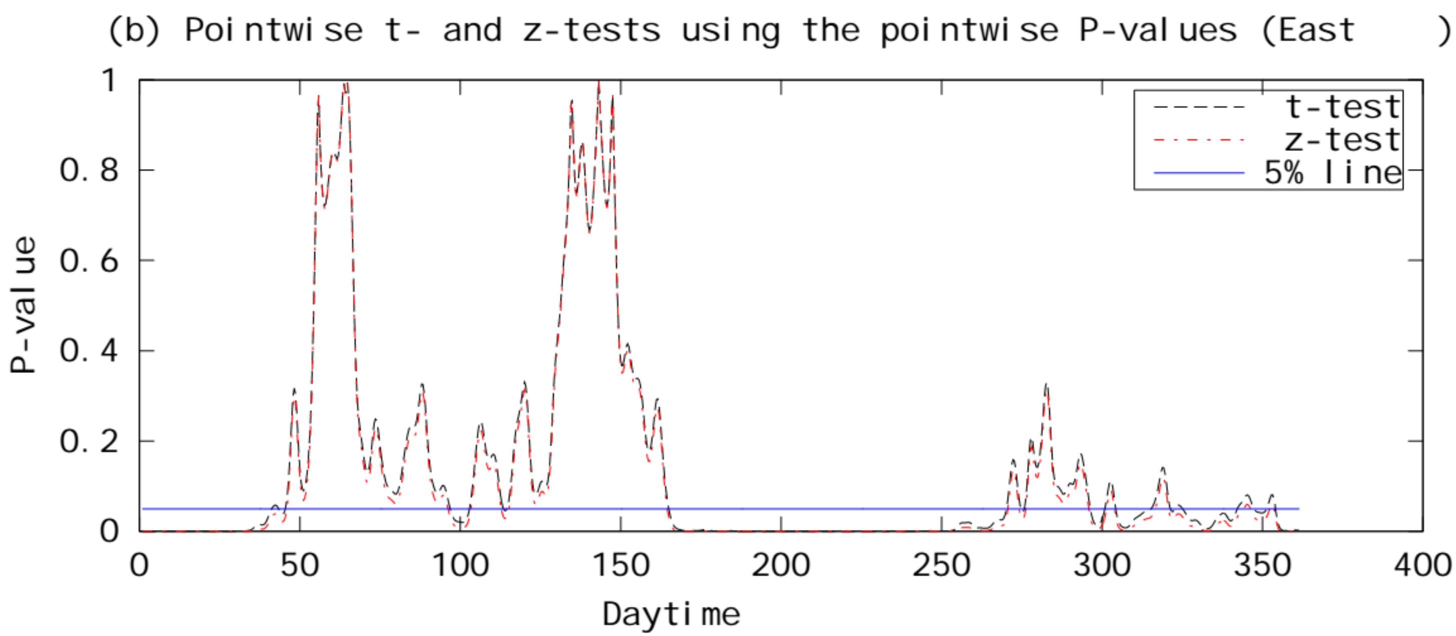


Figure 3
File Edit
Est. cov. func. of the North temperature data
covariance
Daytime
Daytime
A G P R ?
[224.8, 135.6]

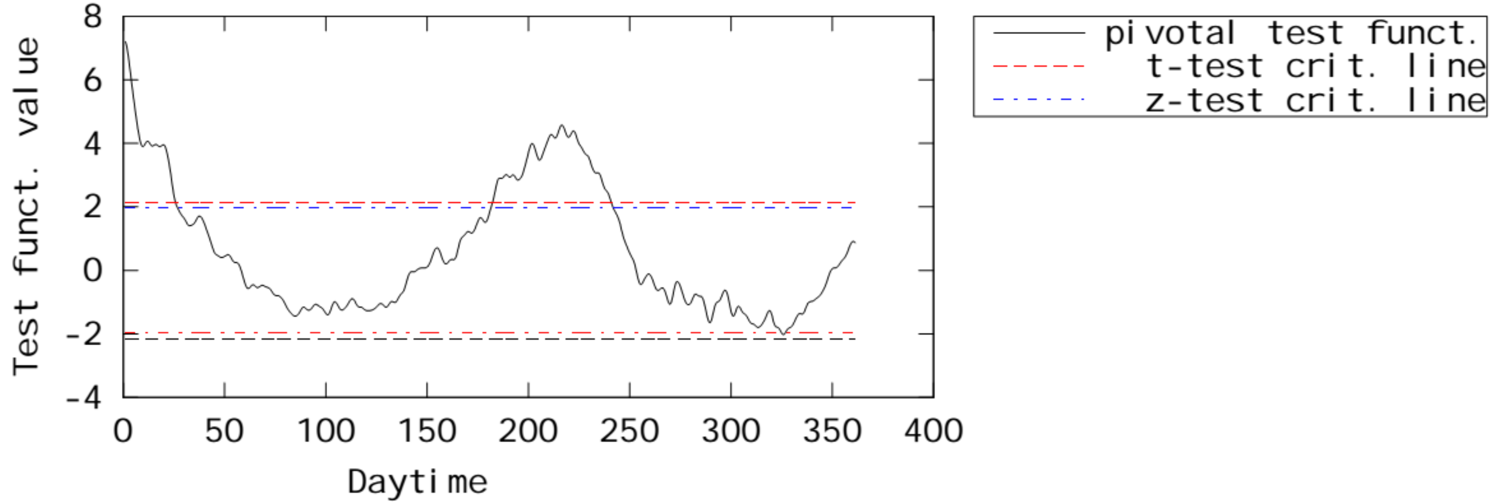


ointwise t- and z-tests using the pivotal test function (East)
Test funct. value
Daytime
pivotal test funct.
t-test crit. line
z-test crit. line
15
10
5
0
-5
0
50
100
150
200
250
300
350
400

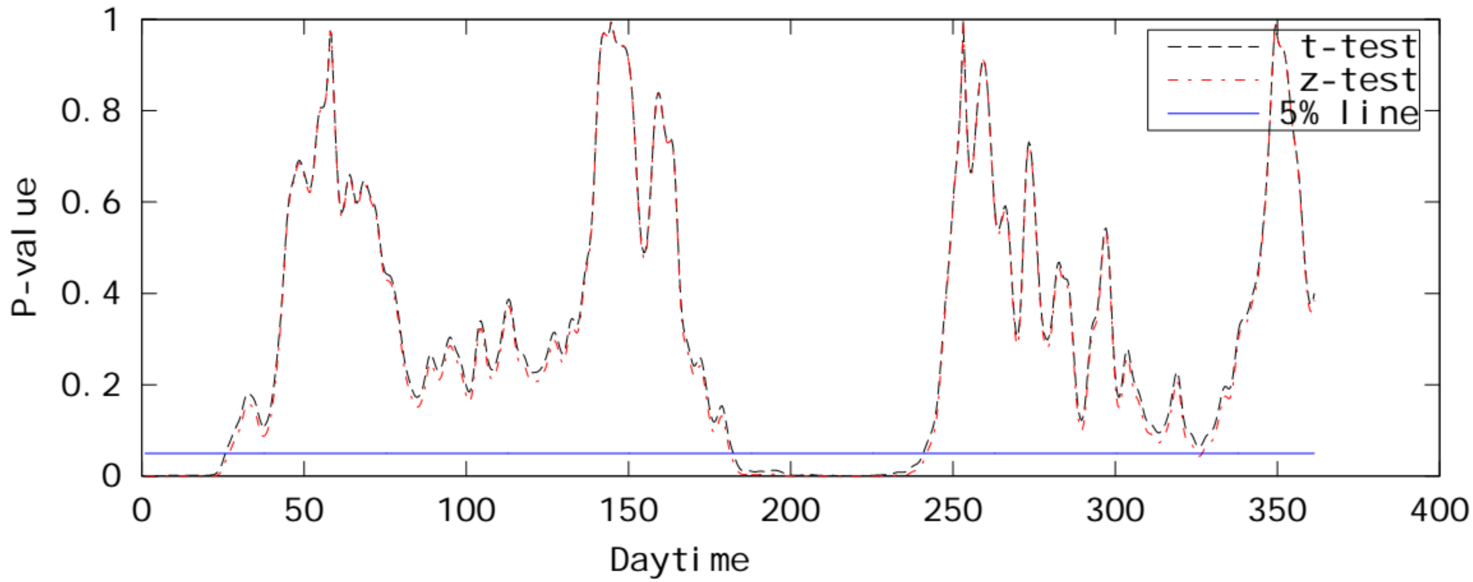


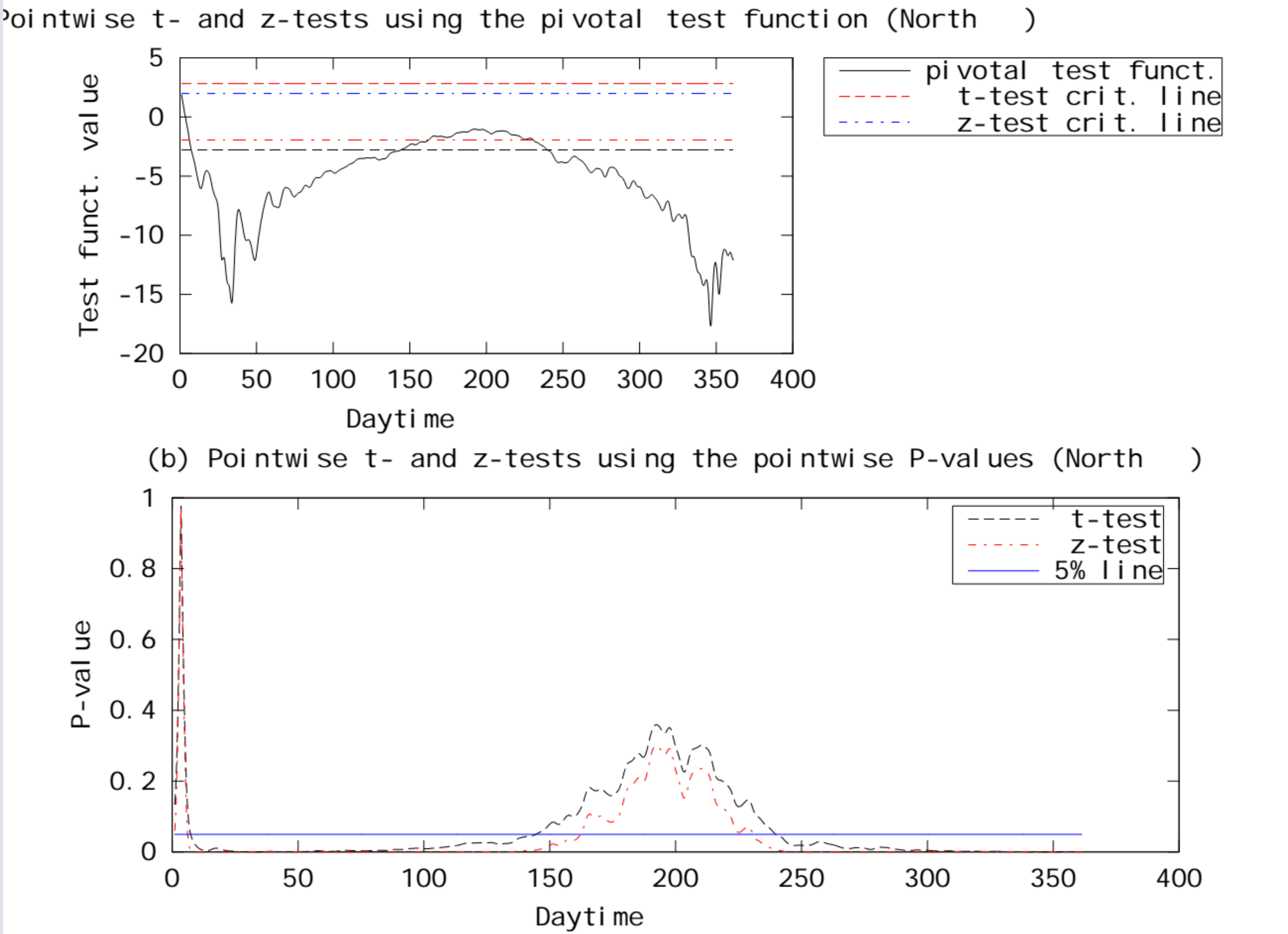
(b) Pointwise t- and z-tests using the pointwise P-values (East)
P-value
Daytime
t-test
z-test
5% line
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
300
350
400

ointwise t- and z-tests using the pivotal test function (West)



(b) Pointwise t- and z-tests using the pointwise P-values (West)




ointwise t- and z-tests using the pivotal test function (North)
Test funct. value
5
0
-5
-10
-15
-20
0 50 100 150 200 250 300 350 400
Daytime
pivotal test funct.
t-test crit. line
z-test crit. line
(b) Pointwise t- and z-tests using the pointwise P-values (North)
P-value
1
0.8
0.6
0.4
0.2
0
0 50 100 150 200 250 300 350 400
Daytime
t-test
z-test
5% line

EXAMPLE

	Method	T_n	$\widehat{\beta}$	$\widehat{d} = \hat{\kappa}$	P-value
East	Naive	301370	8836.1	1.44	1.3901e-008
	Bias-reduced	301370	7467.3	1.57	7.5075e-010
	Bootstrap	301370	---	---	0
West	Naive	175430	2957.6	1.21	0.020
	Bias-reduced	175430	2545.4	1.27	0.013
	Bootstrap	175430	---	---	0.0256
North	Naive	779090	2419.8	1.31	2.7778e-008
	Bias-reduced	779090	1447.3	1.69	1.0874e-012
	Bootstrap	779090	---	---	0

EXAMPLE

	Method	F test	$\hat{\kappa}$	$(n-1)\hat{\kappa}$	P-value
East	Naive	23.61	1.44	20.22	2.1142e-005
	Bias-reduced	23.61	1.57	21.92	1.0749e-005
	Bootstrap	23.61	---	---	0
West	Naive	4.89	1.21	16.99	0.035
	Bias-reduced	4.89	1.27	17.74	0.033
	Bootstrap	4.89	---		0.038
North	Naive	24.61	1.31	5.23	0.003
	Bias-reduced	24.61	1.69	6.75	0.001
	Bootstrap	24.61	---	---	0.33

NEXT WEEK

TWO-SAMPLE PROBLEMS FOR FUNCTIONAL DATA