On the Behrens-Fisher Problem for Functional Data

Jin-Ting Zhang and Xuehua Liang
National University of Singapore

stazjt@nus.edu.sg

International Conference on
Frontiers of Interface between Statistics and Science
30 Dec 2009- 2 Jan 2010, Hyderabad, India
• Behrens-Fisher Problems
• Main Results
• A Simulation Study
• An Application
• Summary
Behrens-Fisher Problems

• BF problem for univariate data
• BF problem for multivariate data
• BF problem for functional data
BF Problem for univariate data

Testing the equality of means of two univariate normal populations with possible unequal variances, i.e., given two samples:

\[ y_{l1}, y_{l2}, \cdots, y_{ln_l} \sim N(\mu_l, \sigma_l^2), l = 1, 2, \]

without assuming \( \sigma_1^2 = \sigma_2^2 \), one wants to test

\[ H_0 : \mu_1 = \mu_2, \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2. \]

When \( \sigma_1^2 = \sigma_2^2 \) is valid, the above problem can be tested using \( t \)-test.
BF Problem for univariate data

BF problem very interesting and drawn attention for decades.

**Existing Methods:**

- Transformation method (Scheffe 1943)
- Approximate degrees of freedom method (Welch 1947)
- Bayes method (Ghosh and Kim 2001)
- Empirical likelihood method (Dong 2004)
**BF Problem for multivariate data**

Testing the equality of mean vectors of two multivariate normal populations with possible unequal covariance matrices, i.e., given two samples:

\[ y_{l1}, y_{l2}, \ldots, y_{ln_l} \sim N_p(\mu_l, \Sigma_l), l = 1, 2, \]

without assuming \( \Sigma_1 = \Sigma_2 \), one wants to test

\[ H_0 : \mu_1 = \mu_2, \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2. \]

When \( \Sigma_1 = \Sigma_2 \) is valid, the above problem can be tested using Hotelling \( T^2 \)-test.
BF Problem for multivariate data

Existing Methods:

- Scheffe’s Transformation method extended to multivariate data by Bennett (1951)
- Scheffe’s Transformation method extended to high-dimensional data by Zhang and Xu (2009)
- Welch’s Approximate degrees of freedom method extended by James (1954), Yao (1965), Johansen (1980), and Nel and van der Merwe (1986)
- Nel and van der Merwe’s test modified and improved by Krishnamoorthy and Yu (2004)
BF Problem for functional data

Testing the equality of mean functions of two Gaussian processes with possible unequal covariance functions, i.e., given two functional samples:

\[ y_{l1}(t), y_{l2}(t), \ldots, y_{ln_l}(t) \sim \text{GP}(\mu_l, \gamma_l), t \in \mathcal{T}, \ l = 1, 2, \]

without assuming \( \gamma_1(s, t) = \gamma_2(s, t) \), one wants to test

\[ H_0 : \mu_1(t) = \mu_2(t), \text{ vs } H_1 : \mu_1(t) \neq \mu_2(t). \]

When \( \gamma_1(s, t) = \gamma_2(s, t) \) is valid, the above problem can be tested using the equal-covariance \( L^2 \)-norm based test (Zhang and Chen 2007) and the equal-covariance \( F \)-type test (Shen and Faraway 2004, Zhang 2009).
Motivating Example: Berkely Growth Data

- Data collected in the Berkeley Growth Study (Tuddenham and Snyder 1954).
- Heights of 54 girls and 39 boys recorded at 31 ages from Year 1 to Year 18.
- Ages not equally spaced.
- Heights of a girl or a boy form a growth curve over ages.
Figure 1: The Berkeley growth data: (a) growth curves for 54 girls; (b) growth curves for 39 boys; (c) mean growth curves and pointwise confidence bands for girls (solid) and boys (dashed); and (d) pointwise standard deviation curves for girls (solid) and boys (dashed).
BF Problem for functional data

Questions:

- Did girls and boys grow in a same pace?
- Did girls and boys grow in a same pace over some growth period?
  - Baby period (Year 1-Year 4)?
  - Post-baby period (Year 4-Year 13)?
  - Teenage period (Year 13-Year 18)?
- Panel (d) suggests that the covariance functions of girls and boys may not be the same.
- Motivating the two-sample BF problem for functional data.
Main Results

• An $F$-type test
• Asymptotic expression
• Null distribution approximation
An $F$-type test

**Test Statistic**

$$
F_n = \frac{\|z\|^2}{\text{tr}(\hat{\gamma}_z)} = \frac{\int_T \left[ \bar{y}_1(t) - \bar{y}_2(t) \right]^2 dt}{\int_T \hat{\gamma}_z(t, t) dt},
$$

- $\text{E}\|z\|^2 = \text{tr}(\gamma_z)$: under $H_0$, $F_n$ is close to 1.
- $n = n_1 + n_2$: total sample size,
- $\bar{y}_l(t), l = 1, 2$: sample mean functions of two samples,
- $z(t) = \bar{y}_1(t) - \bar{y}_2(t) \sim \text{GP}(\mu_z, \gamma_z)$,
- $\mu_z(t) = \mu_1(t) - \mu_2(t)$, $\gamma_z(s, t) = \gamma_1(s, t)/n_1 + \gamma_2(s, t)/n_2$.
- $\hat{\gamma}_z(s, t) = \hat{\gamma}_1(s, t)/n_1 + \hat{\gamma}_2(s, t)/n_2$ where $\hat{\gamma}_l(s, t), l = 1, 2$ are usual sample covariance functions of two samples.
Random expression

Theorem 1  Under $H_0$ and assuming that $\text{tr} (\gamma_l) < \infty, l = 1, 2$, we have

$$F_n \overset{d}{=} \frac{\sum_{r=1}^{\infty} \lambda_r A_r}{\sum_{l=1}^{2} \sum_{r=1}^{\infty} \left[ n_l (n_l - 1) \right]^{-1} \lambda_{lr} A_{lr}},$$

where

- $X \overset{d}{=} Y$ means $X$ and $Y$ have same distribution.
- $A_r \sim \chi^2_{1}$, $A_{lr} \sim \chi^2_{n_l - 1}$, $l = 1, 2$. All of them are independent of each other.
- $\lambda_r$: eigenvalues of $\gamma_z(s,t)$.
- $\lambda_{lr}$: eigenvalues of $\gamma_l(s,t)$, $l = 1, 2$. 
Null distribution approximation

2-cumulant matched $\chi^2$-approx.: Numerator of $F_n$

- Numerator of $F_n$: $S_1 = \sum_{r=1}^{\infty} \lambda_r A_r$, $A_r \sim \chi^1$.
- $S_1$ is a $\chi^2$-type mixture (Zhang 2005).
- Its distribution can be approximated using that of $R_1 = \beta_1 \chi_{d_1}^2$ (Shen and Faraway 2004, Welch 1947).
- Parameters $\beta_1$ and $d_1$ determined via matching 2 cumulants of $S_1$ and $R_1$:

$$\beta_1 = \frac{\sum_{r=1}^{\infty} \lambda_r^2}{\sum_{r=1}^{\infty} \lambda_r} = \frac{\text{tr}(\gamma_z \otimes \gamma_z)}{\text{tr}(\gamma_z)} , \quad d_1 = \frac{(\sum_{r=1}^{\infty} \lambda_r)^2}{\sum_{r=1}^{\infty} \lambda_r^2} = \frac{\text{tr}^2(\gamma_z)}{\text{tr}(\gamma_z \otimes \gamma_z)}.$$

where $\gamma_z \otimes (s, t) = \int_T \gamma_z(s, u) \gamma_z(u, t) du$.
Null distribution approximation

2-cumulant matched $\chi^2$-approx.: Denominator of $F_n$

- Denominator of $F_n$:
  \[ S_2 = \sum_{l=1}^{2} \sum_{r=1}^{\infty} [n_l(n_l - 1)]^{-1}\lambda_{lr}A_{lr}, \quad A_{lr} \sim \chi^2_{n_l-1}, l = 1, 2. \]

- $S_2$ is a $\chi^2$-type mixture (Zhang 2005).

- Its distribution can be approximated using that of $R_2 = \beta_2 \chi^2_{d_2}$.

- Parameters $\beta_2$ and $d_2$ determined via matching 2 cumulants of $S_2$ and $R_2$:
  \[ \beta_2 = \frac{V}{\text{tr}(\gamma_z)}, \quad d_2 = \frac{\text{tr}^2(\gamma_z)}{V}, \]
  where $V = [n_1^2(n_1 - 1)]^{-1}\text{tr}(\gamma_1^{\otimes 2}) + [n_2^2(n_2 - 1)]^{-1}\text{tr}(\gamma_2^{\otimes 2})$. 

Null distribution approximation

2-cumulant matched $\chi^2$-approx.: $F_n \overset{d}{=} \frac{S_1}{S_2}$

\[ F_n \overset{approx.}{\sim} \frac{\beta_1 \chi_{d_1}^2}{\beta_2 \chi_{d_2}^2} = \frac{\chi_{d_1}^2 / d_1}{\chi_{d_2}^2 / d_2} \sim F_{d_1,d_2}, \]

where we use two facts:

- $\beta_1 d_1 = \text{tr}(\gamma_z) = \beta_2 d_2$
- $S_1$ and $S_2$ are independent.
Null distribution approximation

The naive method

- Get sample covariance functions $\hat{\gamma}_l(s, t), l = 1, 2$
- Estimate $\gamma_z(s, t)$ by $\hat{\gamma}_z(s, t) = \hat{\gamma}_1(s, t)/n_1 + \hat{\gamma}_2(s, t)/n_2$
- Estimate $d_1$ and $d_2$ by substituting $\gamma_1(s, t), \gamma_2(s, t)$ and $\gamma_z(s, t)$ in $d_1$ and $d_2$ by $\hat{\gamma}_1(s, t), \hat{\gamma}_2(s, t)$ and $\hat{\gamma}_z(s, t)$ respectively.
- Advantage: Simple and effective
- Drawback: The estimators are biased and can be improved.
A Simulation Study

Testing procedures under consideration:

- ECL$^2$: Equal covariance $L^2$-norm based test (Zhang and Chen 2007)
- ECF: Equal covariance $F$-type test (Shen and Faraway 2004)
- UCF: Unequal covariance $F$-type test proposed in this paper

Aim: Effect of possible unequal covariance functions to ECL$^2$, ECF and UCF
A Simulation Study

Simulation Setup:

\[ y_i(t) = \mu(t) + v_i(t), \ t \in [0, 1], \ i = 1, \ldots, n, \]

\[ \mu(t) = c_1 \cos(c_2 \pi t) + c_3 \sin(c_4 \pi t), \]

\[ v_i(t) = b_{i1} \psi_1(t) + b_{i2} \psi_2(t) + \cdots + b_{i q} \psi_q(t), \]

\[ b_i = [b_{i1}, b_{i2}, \ldots, b_{i q}]^T \sim N_q \left(0, \text{diag}(\lambda_1, \ldots, \lambda_q)\right), \]

- \( c = [c_1, c_2, c_3, c_4]^T \) for \( \mu(t) \) can be flexibly specified
- \( \gamma(s, t) = \sum_{r=1}^{q} \lambda_r \psi_r(s) \psi_r(t), \ s, t \in [0, 1] \) with \( q \) an odd integer
- Set \( \lambda_r = a \rho^r, r = 1, \ldots, q \) for some \( a > 0 \) and \( 0 < \rho < 1 \).
- Set \( \psi_1(t) = 1, \psi_{2r}(t) = \sqrt{2} \sin(2\pi rt), \) and \( \psi_{2r+1}(t) = \sqrt{2} \cos(2\pi rt), \ t \in [0, 1], \ r = 1, \ldots, (q - 1)/2. \)
A Simulation Study

Parameters for two samples: $a_1 = a_2 = 1.3$, $q_1 = q_2 = 13$, and

\[ n_1 = 40, \quad c_1 = [1, 1.5, 2.1, 2.3]^T, \quad \rho_1 = 0.10, \]
\[ n_2 = 60, \quad c_2 = [1 + \Delta, 1.5, 2.1, 2.3]^T, \quad \rho_2 = 0.10 + \Delta \rho, \]

- $\Delta$ specified to control the difference between $\mu_1(t)$ and $\mu_2(t)$
- $\|\mu_1 - \mu_2\|^2 = \Delta^2 / 2$
- $\Delta$ equally spaced in $[0, \Delta_0]$ for some given $\Delta_0$
- $\Delta \rho$ specified to control the difference between $\gamma_1(s, t)$ and $\gamma_2(s, t)$
- $\|\gamma_1 - \gamma_2\|^2 = \sum_{r=1}^{q} [(0.10 + \Delta \rho)^r - 0.10^r]^2$
- $\Delta \rho = 0, 0.25, 0.5$ and $0.75$ specified four cases for consideration
A Simulation Study

- For each pair $(\Delta, \Delta \rho)$, $N = 10000$ pairs of samples generated.
- Each function evaluated at $M = 1000$ equally spaced time points.
- For each pair of samples, test statistics of testing procedures computed.
- For each pair of samples, P-values of testing procedures computed.
- Power = percentage of $\{P-values < .05\}$ in $N$ replications.
Figure 2: *Power functions of ECL$^2$ (crossed curve), ECF (dashed curve) and UCF (solid curve).*
A Simulation Study

Results
From Panel (a): when two covariance functions are equal,

- Powers of the three testing procedures are about the same,
- Type-I errors (powers at $\Delta = 0$) are close to 5%.
- Indicate that application of UCF did not lead to misleading results
A Simulation Study

Results

From Panels (b), (c) and (d): when two covariance functions are different

- Powers of ECL$^2$ and ECF are much lower than those of UCF
- Problem becomes serious with increasing $\Delta \rho$.
- Type-I errors of ECL$^2$ and ECF much lower than 5%
- Type-I errors of UCF are close to 5%.
- Use of ECL$^2$ and ECF may lead to misleading results
A Simulation Study

Simulation Conclusions:

• UCF can be used regardless if two covariance functions are the same

• ECL² and ECF should be avoided unless strong evidence supporting equality of two covariance functions.
An Application

Berkeley Growth Data

• Smoothed using LPK smoothing by Zhang and Chen (2007) to evaluate at any resolution.

• Smoothed separately and using a same bandwidth $h = .3674$ to allow near independence of reconstructions.

• Four growth periods considered: Baby period $[1, 4)$, Post-baby period $[4, 13)$, Teenage period $[13, 18]$ and whole period $[1, 18]$. 
Figure 3: The difference curve (solid) of mean heights of boys subtracting those of girls in the Berkeley growth data, together with its 95% confidence bands (dashed).
Table 1: P-values for testing if boys and girls of the Berkeley growth data have different mean heights over various growth periods (Resolution $M = 1000$).

<table>
<thead>
<tr>
<th>$[a, b]$</th>
<th>ECL$^2$</th>
<th>ECF</th>
<th>UCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1, 4)</td>
<td>$5.64 \times 10^{-3}$</td>
<td>$6.72 \times 10^{-3}$</td>
<td>$6.47 \times 10^{-3}$</td>
</tr>
<tr>
<td>[4, 13)</td>
<td>$3.14 \times 10^{-1}$</td>
<td>$3.16 \times 10^{-1}$</td>
<td>$3.08 \times 10^{-1}$</td>
</tr>
<tr>
<td>[13, 18]</td>
<td>$2.64 \times 10^{-13}$</td>
<td>$7.21 \times 10^{-11}$</td>
<td>$3.80 \times 10^{-10}$</td>
</tr>
<tr>
<td>[1, 18]</td>
<td>$2.23 \times 10^{-7}$</td>
<td>$1.05 \times 10^{-6}$</td>
<td>$1.38 \times 10^{-6}$</td>
</tr>
</tbody>
</table>
Conclusions

- Test results consistent with those observed in Figure 3.
- Over $[1, 4), [4, 13)$ and $[1, 18]$, ECF and UCF have about the same P-values, indicating that effect of possible unequal covariance small.
- Over $[13, 18]$, P-value of UCF more than 5 times of that of ECF, indicating that effect of possible unequal covariance large.
Summary

• UCF proposed and studied
• Its distribution under null hypothesis investigated
• Simulation shows that ECF and ECL² should be avoided if two samples have different covariance functions
• Further study on asymptotic power of the proposed $F$-type test is warranted
• Better methods for estimating $d_1$ and $d_2$ should be developed
THANK YOU