University of St Andrews

Scotland's First University

School of Mathematics and Statistics

Score tests and data-driven tests in directional statistics

Peter Jupp

University of St Andrews

Outline

1. directional statistics
2. testing uniformity
3. hypotheses
4. score tests T_k
5. simpler statistics S_k
6. data-driven versions

1. Directional statistics

Sample space

circle, sphere: S^1, S^{p-1}

projective space: $\mathbb{R}P^{p-1}$

rotation group: $SO(p)$

Stiefel manifold: (frames)

Grassmann manifold: (subspaces)

compact Riemannian manifold M

1. Directional statistics
Sample space

circle, sphere: S^1, S^{p-1}

projective space: $\mathbb{R}P^{p-1}$

rotation group: $SO(p)$

Stiefel manifold: (frames)

Grassmann manifold: (subspaces)

compact Riemannian manifold M

(or a quotient — shape spaces)

Embedding approach

$$\mathbf{t} : M \to L^2(M)$$

vector space

Embedding approach

$$\mathbf{t}_k : M \to E_k \subset L^2(M)$$

E_k are finite-dimensional, orthogonal

E_k is kth eigenspace of Laplacian $\qquad k = 1, 2, \ldots$

For S^1,

$$E_k = \operatorname{span}\{\cos(k\theta), \sin(k\theta)\}$$

$$\mathbf{t}_k(x) = \cos(k(\theta - x))$$

2. Testing uniformity

Given $a_1, a_2, \ldots$,

$$\mathbf{t}(x) = \sum_{k=1}^{\infty} a_k \mathbf{t}_k(x) \qquad \in L^2(M)$$

Summarise sample $x_1, \ldots, x_n$ by

$$\bar{\mathbf{t}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{t}(x_i) \qquad \in L^2(M)$$

Reject uniformity if $\|\bar{\mathbf{t}}\|^2$ large.

E.g. S^1

$$\bar{R}_k = \left\| \frac{1}{n} \sum_{i=1}^{n} (\cos kx_i, \sin kx_i) \right\|$$

Reject uniformity if

$$\sum_{k=1}^{\infty} {a_k}^2 \bar{R}_k^2 \qquad \text{large}$$

generalised Rayleigh test

Problem

How to choose $a_1, a_2, \ldots$?

$$\text{few } a_k \neq 0 \quad \Rightarrow \quad \text{(often) simple to calculate}$$

$$\text{all } a_k \neq 0 \quad \Rightarrow \quad \text{consistent against all alternatives}$$

Embarrassment of choice!

Score tests of uniformity

$$(a_1, a_2, \ldots) = (\underbrace{1, 1, \ldots, 1}_{k}, 0, 0, \ldots)$$

$$\mathbf{t}_{(k)}(x) = (\mathbf{t}_1(x), \ldots, \mathbf{t}_k(x)) \in \oplus_{j=1}^{k} E_j \subset L^2(M)$$

$$\bar{\mathbf{t}}_{(k)} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{t}_{(k)}(x_i)$$

$$T_k = n \|\bar{\mathbf{t}}_{(k)}\|^2$$

Reject uniformity if

T_k large

Select k by BIC

$\hat{k}$ is value of k maximising

$$T_k - \nu_k c(n)$$

$\nu_k = \sum_{j=1}^{k} \dim E_j$ c increasing, e.g. $c(n) = \log n$

Reject uniformity for $T_{\hat{k}}$ large

'Penalised score test' —

penalises higher-dimensional models

Nice properties

$$P(\hat{k} = \infty) = 0$$

Under uniformity, as $n \to \infty$,

$$\begin{aligned} \hat{k} &\to 1 \\ T_{\hat{k}} &\dot{\sim} \chi^2_{\nu_1} \end{aligned}$$

Consistent against all alternatives

3. Hypotheses

uniformity

q populations equal

symmetry group G acts on M

e.g. $M = S^2$, $G = SO(2)$, rotation about axis

independence X on M, Y on N

goodness of fit $f(\cdot;\theta)$

4. Score tests T_k

H_0	T_k
uniformity	$n\lVert\bar{\mathbf{t}}_{(k)}\rVert^2$
symmetry	$n\,\bar{\mathbf{t}}_{(k)-}{}^{\mathrm{T}}\mathbf{S}_{(k)-}{}^{-1}\bar{\mathbf{t}}_{(k)-}$
q populations	$\dfrac{n-q}{q-1}\,\mathrm{tr}\left(\mathbf{B}_{(k)}\mathbf{W}_{(k)}{}^{-1}\right)$
independence	$nr^2_{(k)}$
goodness of fit	$n\,\bar{\mathbf{t}}_{(k)w}{}^{\mathrm{T}}\mathbf{S}_{(k)w}{}^{-1}\bar{\mathbf{t}}_{(k)w}$

symmetry

Group G acts on M, and so on each E_k

$$
\begin{aligned}
E_{k+} &= \{f \in E_k : f(gx) = f(x), g \in G\} \\
E_{k-} &= \left\{f \in E_k : \int_G f(gx)dg = 0\right\}
\end{aligned}
$$

$$
\begin{aligned}
E_k &= E_{k+} \oplus E_{k-} \\
\mathbf{t}_{(k)} &= (\mathbf{t}_{(k)+}, \mathbf{t}_{(k)-})
\end{aligned}
$$

$$
\bar{\mathbf{t}}_{(k)-} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{t}_{(k)-}(x_i)
$$

score tests T_k

H_0	T_k
uniformity	$n\|\bar{\mathbf{t}}_{(k)}\|^2$
symmetry	$n\,\bar{\mathbf{t}}_{(k)-}{}^{\mathrm{T}}\mathbf{S}_{(k)-}{}^{-1}\bar{\mathbf{t}}_{(k)-}$
q populations	$\frac{n-q}{q-1}\,\mathrm{tr}\left(\mathbf{B}_{(k)}\mathbf{W}_{(k)}{}^{-1}\right)$
independence	$nr^2_{(k)}$
goodness of fit	$n\,\bar{\mathbf{t}}_{(k)w}{}^{\mathrm{T}}\mathbf{S}_{(k)w}{}^{-1}\bar{\mathbf{t}}_{(k)w}$

q populations

$$(x_{11}, \ldots, x_{1n_1}), \ldots, (x_{q1}, \ldots, x_{qn_q})$$

$$\mathbf{B}_{(k)} = \sum_{i=1}^{q} \sum_{j=1}^{n_i} \left(\mathbf{t}_{(k)}(x_{ij}) - \bar{\mathbf{t}}_{(k)i\cdot}\right) \left(\mathbf{t}_{(k)}(x_{ij}) - \bar{\mathbf{t}}_{(k)i\cdot}\right)^{\mathrm{T}}$$

$$\mathbf{W}_{(k)} = \sum_{i=1}^{q} n_i \left(\bar{\mathbf{t}}_{(k)i\cdot} - \bar{\mathbf{t}}_{(k)\cdot\cdot}\right) \left(\bar{\mathbf{t}}_{(k)i\cdot} - \bar{\mathbf{t}}_{(k)\cdot\cdot}\right)^{\mathrm{T}}$$

score tests T_k

H_0	T_k
uniformity	$n\|\bar{\mathbf{t}}_{(k)}\|^2$
symmetry	$n\,\bar{\mathbf{t}}_{(k)-}{}^{\mathrm{T}}\mathbf{S}_{(k)-}{}^{-1}\bar{\mathbf{t}}_{(k)-}$
q populations	$\frac{n-q}{q-1}\,\mathrm{tr}\left(\mathbf{B}_{(k)}\mathbf{W}_{(k)}{}^{-1}\right)$
independence	$nr^2_{(k)}$
goodness of fit	$n\,\bar{\mathbf{t}}_{(k)w}{}^{\mathrm{T}}\mathbf{S}_{(k)w}{}^{-1}\bar{\mathbf{t}}_{(k)w}$

independence

$$\mathbf{t}_{(k)} : M \to L^2(M) \qquad \mathbf{u}_{(k)} : N \to L^2(N)$$

$$\mathrm{var}\begin{pmatrix} \mathbf{t}_{(k)}(x) \\ \mathbf{u}_{(k)}(y) \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{(k)11} & \mathbf{S}_{(k)12} \\ \mathbf{S}_{(k)21} & \mathbf{S}_{(k)22} \end{pmatrix}$$

$$r^2_{(k)} = \mathrm{tr}\left(\mathbf{S}_{(k)11}{}^{-1}\mathbf{S}_{(k)12}\mathbf{S}_{(k)22}{}^{-1}\mathbf{S}_{(k)21}\right)$$

score tests T_k

H_0	T_k
uniformity	$n\|\bar{\mathbf{t}}_{(k)}\|^2$
symmetry	$n\,\bar{\mathbf{t}}_{(k)-}{}^{\mathrm{T}}\mathbf{S}_{(k)-}{}^{-1}\bar{\mathbf{t}}_{(k)-}$
q populations	$\frac{n-q}{q-1}\,\mathrm{tr}\left(\mathbf{B}_{(k)}\mathbf{W}_{(k)}{}^{-1}\right)$
independence	$nr^2_{(k)}$
goodness of fit	$n\,\bar{\mathbf{t}}_{(k)w}{}^{\mathrm{T}}\mathbf{S}_{(k)w}{}^{-1}\bar{\mathbf{t}}_{(k)w}$

goodness of fit

$$\mathbf{t}_{(k)w}(x) = \frac{1}{f(x;\hat{\theta})}\mathbf{t}_{(k)}(x)$$

$$\bar{\mathbf{t}}_{(k)w} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{t}_{(k)w}(x_i)$$

Aside: cubic corrections

Under H_0,

$$T_k \ \dot{\sim} \ \chi^2_{\nu_k} \qquad\qquad \text{error } O(n^{-1})$$

$$T_k^* = \left\{ 1 + \frac{1}{n} \left[c_0 + c_1 T_k + c_2 T_k^2 \right] \right\} T_k$$

Under H_0,

$$T_k^* \ \dot{\sim} \ \chi^2_{\nu_k} \qquad\qquad \text{error } O(n^{-2})$$

5. Simpler statistics S_k

H_0	S_k
uniformity	$n\|\bar{\mathbf{t}}_{(k)}\|^2$
symmetry	$n\|\bar{\mathbf{t}}_{(k)-}\|^2$
q populations	$\mathrm{tr}\left(\mathbf{B}_{(k)}\right)$
independence	$\mathrm{tr}\left(\mathbf{S}_{12(k)}\mathbf{S}_{21(k)}\right)$
goodness of fit	$\|\bar{\mathbf{t}}_{(k)w}\|^2$

6. Data-driven versions

Select k by BIC

$\hat{k}$ is value of k maximising $T_k - \nu_k c(n)$

or

$\hat{k}$ is value of k maximising $S_k - \hat{\nu}_k c(n)$

Data-driven versions of T_k

$1 \leq k \leq K$

$\hat{k}$ is value of k maximising $T_k - \nu_k c(n)$

H_0	T_k	ν_k
symmetry	$n\,\bar{\mathbf{t}}_{(k)-}{}^{\mathrm{T}}\mathbf{S}_{(k)-}{}^{-1}\bar{\mathbf{t}}_{(k)-}$	ν_{k-}
q pop'ns	$\frac{n-q}{q-1}\,\mathrm{tr}\left(\mathbf{B}_{(k)}\mathbf{W}_{(k)}{}^{-1}\right)$	$(q-1)\nu_k$
indep'ce	$nr^2_{(k)}$	$\sum_{r+s=k} \nu_{M,r}\nu_{N,s}$
g.o.f.	$n\,\bar{\mathbf{t}}_{(k)w}{}^{\mathrm{T}}\mathbf{S}_{(k)w}{}^{-1}\bar{\mathbf{t}}_{(k)w}$	ν_k

Nice properties of $T_{\hat{k}}$

Under H_0, as $n \to \infty$,

$$\begin{aligned} \hat{k} &\to 1 \\ T_{\hat{k}} &\dot{\sim} \chi^2_{\nu_1} \end{aligned}$$

Consistent against alternatives with

$$E\left[\mathbf{t}_{(K)}(x)\right] \neq \mathbf{0}.$$

Data-driven versions of S_k

$$1 \leq k \leq K$$

$\hat{k}$ is value of k maximising $S_k - \hat{\nu}_k c(n)$

H_0	S_k	$\hat{\nu}_k$
symmetry	$n\|\bar{\mathbf{t}}_{(k)-}\|^2$	$\mathrm{tr}\left(\mathbf{S}_{(k)-}\right)$
q populations	$\mathrm{tr}\left(\mathbf{B}_{(k)}\right)$	$\frac{q-1}{n-q}\mathrm{tr}\left(\mathbf{W}_{(k)}\right)$
independence	$\mathrm{tr}\left(\mathbf{S}_{12(k)}\mathbf{S}_{21(k)}\right)$	$\mathrm{tr}\left(\mathbf{S}_{11(k)}\right)\mathrm{tr}\left(\mathbf{S}_{22(k)}\right)$
goodness of fit	$\|\bar{\mathbf{t}}_{(k)w}\|^2$	$\frac{1}{n}\sum_{i=1}^{n}\frac{\|\mathbf{t}_{(k)}(x_i)\|^2}{f(x_i;\hat{\theta})}$

Nice properties of $S_{\hat{k}}$

Under H_0, as $n \to \infty$,

$$\begin{aligned} \hat{k} &\to 1 \\ S_{\hat{k}} &\dot{\sim} \text{ non-degenerate distribution} \end{aligned}$$

Consistent against alternatives with

$$E\left[\mathbf{t}_{(K)}(x)\right] \neq \mathbf{0}.$$