Kryštof Eben

Classification into two von Mises distributions with unknown mean directions

Aplikace matematiky, Vol. 28 (1983), No. 3, 230–237

Persistent URL: `http://dml.cz/dmlcz/104029`

CLASSIFICATION INTO TWO VON MISES DISTRIBUTIONS WITH UNKNOWN MEAN DIRECTIONS

Kryštof Eben

(Received November 26, 1982)

Classification procedures for directional data have been studied by Morris and Laycock [3] for the case of two von Mises distributions on the circle and two Fisher distributions on the sphere, both with known parameters. The present paper deals with two von Mises distributions with unknown mean directions and a common concentration parameter that is known. The likelihood rule and the plug-in rule are examined. For the statistic of the plug-in rule, the moment generating function is given. The moments are obtained as well but the higher ones are too complicated for practical use.

1. FORMULATION OF THE PROBLEM AND RELATED CLASSIFICATION RULES

In the text we use the modified Bessel functions of the first kind

$$I_p(z) = \frac{1}{2\pi}\int_0^{2\pi} \cos p\theta \exp\{z\cos\theta\}\, d\theta = \sum_{\nu=0}^{\infty} \frac{(z/2)^{2\nu+p}}{\nu!\,\Gamma(\nu+p+1)}.$$

We need it only for whole orders $p \in Z$ and z real, although it is defined more generally. Further, in our notation we shall not distinguish random variables from the arguments of the corresponding densities.

Let $\boldsymbol{l} = (\cos\theta, \sin\theta)'$, $0 \leqq \theta < 2\pi$ be a random unit vector in the plane taking values on the unit circle. The density of $\boldsymbol{l}$ means the density with respect to the Lebesgue measure on the circle. Transforming the density of $\boldsymbol{l}$ to polar coordinates we obtain the density (with respect to the ordinary Lebesgue measure) of the random angle θ.

$\boldsymbol{l}$ has a von Mises distribution $M(\boldsymbol{m}, \varkappa)$ with mean direction $\boldsymbol{m} = (\cos\mu, \sin\mu)'$ and concentration parameter $\varkappa > 0$, if the density of $\boldsymbol{l}$ is given by

$$f(\boldsymbol{l};\, \boldsymbol{m}, \varkappa) = c(\varkappa)\exp\{\varkappa \boldsymbol{m}'\boldsymbol{l}\}\,, \quad \varkappa > 0\,, \quad \boldsymbol{m}'\boldsymbol{m} = \boldsymbol{l}'\boldsymbol{l} = 1\,. \tag{1.1}$$

Using the polar coordinates, we obtain this expression in the form of

$$(1.2) \qquad g(\theta; \mu, \varkappa) = c(\varkappa) \exp\{\varkappa \cos(\theta - \mu)\}, \quad \theta, \mu \in \langle 0, 2\pi), $$

$c(\varkappa) = 1/2\pi I_0(\varkappa)$ being the norming constant.

Consider the problem of classification into two von Mises distributions $M(\boldsymbol{m}_1, \varkappa)$, $M(\boldsymbol{m}_2, \varkappa)$, where $\varkappa$ is known, $\boldsymbol{m}_1 = (\cos \mu_1, \sin \mu_1)'$, $\boldsymbol{m}_2 = (\cos \mu_2, \sin \mu_2)'$. If $\boldsymbol{m}_1, \boldsymbol{m}_2$ were known, the Bayes discriminant function would be as derived in [3]:

$$(1.3) \qquad U^* = \cos(\theta - \mu_1) - \cos(\theta - \mu_2).$$

Assume that $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ are unknown. Let $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$ and $\boldsymbol{l}_{r+1}, \ldots, \boldsymbol{l}_n$ be independent random samples from $M(\boldsymbol{m}_1, \varkappa)$ and $M(\boldsymbol{m}_2, \varkappa)$. If we adopt the likelihood approach to classification, the problem may be viewed as that of testing the hypothesis ($\boldsymbol{l}$ denotes the classified vector)

$$H_0 : \boldsymbol{l}_1, \ldots, \boldsymbol{l}_r, \boldsymbol{l} \sim M(\boldsymbol{m}_1, \varkappa)$$
$$\boldsymbol{l}_{r+1}, \ldots, \boldsymbol{l}_n \sim M(\boldsymbol{m}_2, \varkappa)$$

against the alternative

$$A \ : \boldsymbol{l}_1, \ldots, \boldsymbol{l}_r \sim M(\boldsymbol{m}_1, \varkappa)$$
$$\boldsymbol{l}, \boldsymbol{l}_{r+1}, \ldots, \boldsymbol{l}_n \sim M(\boldsymbol{m}_2, \varkappa).$$

The usual likelihood statistic for this kind of testing equals in our case

$$\lambda = \frac{\max\limits_{\boldsymbol{m}_1, \boldsymbol{m}_2} [c(\varkappa)]^{n+1} \exp\{\varkappa(\sum\limits_{i=1}^{r} \boldsymbol{l}_i' \boldsymbol{m}_1 + (\sum\limits_{i=r+1}^{n} \boldsymbol{l}_i' + \boldsymbol{l}') \boldsymbol{m}_2)\}}{\max\limits_{\boldsymbol{m}_1, \boldsymbol{m}_2} [c(\varkappa)]^{n+1} \exp\{\varkappa((\sum\limits_{i=1}^{r} \boldsymbol{l}_i' + \boldsymbol{l}') \boldsymbol{m}_1 + \sum\limits_{i=r+1}^{n} \boldsymbol{l}_i' \boldsymbol{m}_2)\}}.$$

In both the upper and the lower term of this ratio the maximum over all unit vectors $\boldsymbol{m}_1, \boldsymbol{m}_2$ is achieved for $\boldsymbol{m}_1, \boldsymbol{m}_2$ maximizing the corresponding scalar products; i.e., in the numerator for

$$\boldsymbol{m}_1^* = (\sum_{i=1}^{r} \boldsymbol{l}_i)/R_1, \qquad R_1 = \|\sum_{i=1}^{r} \boldsymbol{l}_i\|, \ ^{*)}$$

$$\tilde{\boldsymbol{m}}_2 = (\sum_{i=r+1}^{n} \boldsymbol{l}_i + \boldsymbol{l})/R_{2,0}, \quad R_{2,0} = \|\sum_{i=r+1}^{n} \boldsymbol{l}_i + \boldsymbol{l}\|,$$

and in the denominator for

$$\tilde{\boldsymbol{m}}_1 = (\sum_{i=1}^{r} \boldsymbol{l}_i + \boldsymbol{l})/R_{1,0}, \quad R_{1,0} = \|\sum_{i=1}^{r} \boldsymbol{l}_i + \boldsymbol{l}\|,$$

$$\boldsymbol{m}_2^* = (\sum_{i=r+1}^{n} \boldsymbol{l}_i)/R_2, \qquad R_2 = \|\sum_{i=r+1}^{n} \boldsymbol{l}_i\|.$$

---

*) We use to exclude the event $\sum \boldsymbol{l}_i = 0$ which is of measure zero.

Thus we obtain

$$\lambda = \exp\{\varkappa(R_1 + R_{2,0} - R_{1,0} - R_2)\}\,,$$

which is equivalent to the statistic $R_1 + R_{2,0} - R_{1,0} - R_2$.

Remark. The vectors $\boldsymbol{m}_1^*$ and $\boldsymbol{m}_2^*$ are the sample mean directions. They are maximum likelihood estimates of $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ based on the samples $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$ and $\boldsymbol{l}_{r+1}, \ldots, \boldsymbol{l}_n$.

Let us denote $\boldsymbol{l}_i = (\cos\theta_i, \sin\theta_i)'$, $i = 1, \ldots, n$, $\theta_i \in \langle 0, 2\pi)$, $\boldsymbol{m}_1^* = (\cos\psi_1, \sin\psi_1)'$, $\boldsymbol{m}_2^* = (\cos\psi_2, \sin\psi_2)'$, $\psi_1, \psi_2 \in \langle 0, 2\pi)$. Then

$$(1.4)\quad R_1 - R_{1,0} = R_1 - ((R_1\cos\psi_1 + \cos\theta)^2 + (R_1\sin\psi_1 + \sin\theta)^2)^{1/2} =$$
$$= (R_1^2)^{1/2} - (R_1^2 + 2R_1\cos(\psi_1 - \theta) + 1)^{1/2} =$$
$$= \frac{-2R_1\cos(\psi_1 - \theta) - 1}{(R_1^2)^{1/2} + (R_1^2 + 2R_1\cos(\psi_1 - \theta) + 1)^{1/2}} =$$
$$= \frac{-2\cos(\psi_1 - \theta) - 1/R_1}{1 + (1 + (2\cos(\psi_1 - \theta))/R_1 + 1/R_1^2)^{1/2}}\,.$$

The distribution of R_1 does not depend on $\boldsymbol{m}$ (see [2], § 4.5) and it is the same as if $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$ came from $M((1, 0)', \varkappa)$. In that case

$$\mathsf{E}\cos\theta_i = \frac{1}{2\pi I_0(\varkappa)}\int_0^{2\pi}\cos\theta\, e^{\varkappa\cos\theta}\,d\theta = \frac{I_1(\varkappa)}{I_0(\varkappa)} > 0$$

and according to the strong law of large numbers

$$\lim_{r\to\infty}\sum_{i=1}^{r}\cos\theta_i = \infty \quad \text{a.s.}$$

As $R_1 \geqq \sum_{i=1}^{r}\cos\theta_i$, we have $\lim_{r\to\infty} R_1 = \infty$ a.s. Returning back to (1.4), we obtain

$$\lim_{r\to\infty}(R_1 - R_{1,0}) = -\cos(\psi_1 - \theta)\,.$$

Similarly we state

$$\lim_{n-r\to\infty}(R_2 - R_{2,0}) = -\cos(\psi_2 - \theta)\,.$$

Thus the likelihood statistic is asymptotically equivalent to the statistic $-U = \cos(\theta - \psi_2) - \cos(\theta - \psi_1)$. The statistic $U = \cos(\theta - \psi_1) - \cos(\theta - \psi_2)$ is immediately seen to be the statistic of the plug-in rule derived from (1.3) by substituting the estimates ψ_1, ψ_2 instead of the true parameters μ_1, μ_2.

2. THE MOMENT GENERATING FUNCTION AND MOMENTS OF THE STATISTIC U

Our starting point will be the following theorem (for the proof see [2], § 4.5):

Theorem. *Let* $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$ *be a random sample from* $M(\boldsymbol{m}, \varkappa)$, $\boldsymbol{m} = (\cos \mu, \sin \mu)'$, $\sum_{i=1}^{r} \boldsymbol{l}_i = R(\cos \psi, \sin \psi)' = R\boldsymbol{m}^{*\prime}$. *Then*

(i) *the distribution of* R *does not depend on* μ *(i.e.* R *is an ancillary statistic for* μ*)*

(ii) *the conditional distribution of* $\boldsymbol{m}^*$ *given* R *is* $M(\boldsymbol{m}, \varkappa R)$.

Motivated by this theorem and the Fisher ancillary principle we shall assume

$$\boldsymbol{m}_1^* \sim M(\boldsymbol{m}_1, \varkappa R_1)\,, \quad \boldsymbol{m}_2^* \sim M(\boldsymbol{m}_2, \varkappa R_2)\,.$$

We compute the moment generating function of U for $\boldsymbol{l} \sim M(\boldsymbol{m}_1, \varkappa)$, the other case is symmetric. Taking (1.2) for the densities,

$$(2.1)\quad \mathsf{E}\, e^{tU} = c(\varkappa)\, c(\varkappa R_1)\, c(\varkappa R_2) \int_0^{2\pi}\!\!\int_0^{2\pi}\!\!\int_0^{2\pi} \exp\{t(\cos(\theta - \psi_1) - \cos(\theta - \psi_2))\} \times$$

$$\times \exp\{\varkappa(R_1 \cos(\psi_1 - \mu_1) + R_2 \cos(\psi_2 - \mu_2) + \cos(\theta - \mu_1))\}\, d\psi_1\, d\psi_2\, d\theta =$$

$$= c(\varkappa)\, c(\varkappa R_1)\, c(\varkappa R_2) \int_0^{2\pi} \left[\int_0^{2\pi} \exp\{t \cos(\theta - \psi_1) + \varkappa R_1 \cos(\psi_1 - \mu_1)\}\, d\psi_1\right] \times$$

$$\times \left[\int_0^{2\pi} \exp\{-t \cos(\theta - \psi_2) + \varkappa R_2 \cos(\psi_2 - \mu_2)\}\, d\psi_2\right] \times$$

$$\times \exp\{\varkappa \cos(\theta - \mu_1)\}\, d\theta\,.$$

The first bracket in (2.1) equals to

$$(2.2)\qquad \int_0^{2\pi} \exp\{a \cos \psi_1 + b \sin \psi_1\}\, d\psi_1\,,$$

where

$$a = t \cos \theta + \varkappa R_1 \cos \mu_1\,, \quad b = t \sin \theta + \varkappa R_1 \sin \mu_1\,.$$

Setting

$$\varrho = (a^2 + b^2)^{1/2}\,, \quad a = \varrho \cos \varphi\,, \quad b = \varrho \sin \varphi\,,$$

(2.2) becomes

$$\int_0^{2\pi} \exp\{\varrho \cos(\psi_1 - \varphi)\}\, d\psi_1 = 2\pi I_0(\varrho) =$$

$$= 2\pi I_0((t^2 + \varkappa^2 R_1^2 + 2t\varkappa R_1 \cos(\theta - \mu_1))^{1/2})\,.$$

By the same argument, the second bracket in (2.1) equals to

$$2\pi I_0(\tilde{\varrho}) = 2\pi I_0((t^2 + \varkappa^2 R_2^2 + 2t\varkappa R_2 \cos(\theta - \mu_2 - \pi))^{1/2})$$

and (2.1) simplifies to

$$\mathsf{E}\, e^{tU} = [2\pi I_0(\varkappa) I_0(\varkappa R_1) I_0(\varkappa R_2)]^{-1} \int_0^{2\pi} I_0(\varrho) I_0(\tilde{\varrho}) \exp\{\varkappa \cos(\theta - \mu_1)\} \, d\theta \,. \tag{2.3}$$

Lemma. (a) *The von Neumann summation formula. For x_1, x_2, θ real,*

$$I_0((x_1^2 + x_2^2 + 2x_1x_2\cos\theta)^{1/2}) = \tag{2.4}$$
$$= I_0(x_1) I_0(x_2) + 2\sum_{p=1}^{\infty} I_p(x_1) I_p(x_2) \cos p\theta \,.$$

(b) *The sum on the right-hand side of (2.4) is absolutely convergent.*
(c) *For every $p \in Z$, $\varkappa > 0$,*

$$\int_0^{2\pi} \sin p\theta \exp\{\varkappa \cos\theta\} \, d\theta = 0 \,.$$

(d) *For $p, q \in Z$, μ_1, μ_2 real, $\varkappa > 0$,*

$$\int_0^{2\pi} \cos p(\theta - \mu_1) \cos q(\theta - \mu_2 - \pi) \exp\{\varkappa \cos(\theta - \mu_1)\} \, d\theta = \frac{(-1)^q}{2} \times$$
$$\times \int_0^{2\pi} [\cos((p+q)(\theta - \mu_1) + q(\mu_1 - \mu_2)) + \cos((p-q)(\theta - \mu_1) - q(\mu_1 - \mu_2))] \times$$
$$\times \exp\{\varkappa \cos(\theta - \mu_1)\} \, d\theta = \frac{(-1)^q}{2} \cos q(\mu_1 - \mu_2) \, 2\pi(I_{p+q}(\varkappa) + I_{p-q}(\varkappa)) \,.$$

For the derivation of (2.4) see e.g. [1] § 7·6. The absolute convergence may be obtained for example from the rough inequality

$$|I_p(x)| \leqq \sum_{\nu=0}^{\infty} \frac{|x/2|^{2\nu+p}}{\nu!\,(\nu+p)!} \leqq \frac{|x/2|^p}{p!} \exp(x/2)^2 \,.$$

(c) is self-evident and (d) may be easily checked with the help of (c).

The above lemma enables us to expand the term $I_0(\varrho) I_0(\tilde{\varrho})$ in the integrand of (2.3):

$$I_0(\varrho) I_0(\tilde{\varrho}) = I_0^2(t) I_0(\varkappa R_1) I_0(\varkappa R_2) +$$
$$+ 2 I_0(t) I_0(\varkappa R_1) \sum_{q=1}^{\infty} I_q(t) I_q(\varkappa R_2) \cos q(\theta - \mu_2 - \pi) +$$
$$+ 2 I_0(t) I_0(\varkappa R_2) \sum_{p=1}^{\infty} I_p(t) I_p(\varkappa R_1) \cos p(\theta - \mu_1) +$$
$$+ 4 \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} I_p(t) I_q(t) I_p(\varkappa R_1) I_q(\varkappa R_2) \cos p(\theta - \mu_1) \cos q(\theta - \mu_2 - \pi) \,.$$

Since the series involved are absolutely convergent, we can integrate termwise. The calculation needs only (c) and (d) of the lemma and is omitted here. As a result we obtained the following expression for the moment generating function of the statistic U:

$$\mathsf{E}\, e^{tU} = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} a_{pq} I_p(t) I_q(t),$$

where the coefficients a_{pq} are defined in subsequent manner: let us denote $\alpha_p(\varkappa) = I_p(\varkappa)/I_0(\varkappa)$, $\delta = |\mu_1 - \mu_2|$; then

$$\begin{aligned} a_{00} &= 1, \\ a_{0q} &= (-1)^q\, 2\, \alpha_q(\varkappa)\, \alpha_q(\varkappa R_2) \cos q\delta, \\ a_{p0} &= 2\, \alpha_p(\varkappa)\, \alpha_p(\varkappa R_1), \\ a_{pq} &= (-1)^q\, 2\, \alpha_p(\varkappa R_1)\, \alpha_q(\varkappa R_2)\, (\alpha_{p+q}(\varkappa) + \alpha_{p-q}(\varkappa)) \cos q\delta, \\ & p, q \geqq 1. \end{aligned}$$

3. THE MOMENTS OF THE STATISTIC U

For the functions $I_p(t)$ we have (see [1] 7.2.5, resp. §7.11)

$$I_p(0) = 0 \quad \text{for} \quad p \neq 0, \qquad = 1 \quad \text{for} \quad p = 0, \tag{3.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} I_p(t) = \tfrac{1}{2}(I_{p-1}(t) + I_{p+1}(t)). \tag{3.2}$$

Let us consider derivatives of the term $I_p(t)\, I_q(t)$. According to (3.1) and (3.2),

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} I_p(t)\, I_q(t)\Big|_{t=0} = c_{pq}^{(k)},$$

where $c_{pq}^{(k)}$ is the coefficient standing by $I_0(t)\, I_0(t)$ in the expression for $(\mathrm{d}^k/\mathrm{d}t^k) I_p(t) I_q(t)$. Obviously $c_{pq}^{(k)} = 0$ for $p + q > k$ and the k-th moment of U equals

$$\mathsf{E} U^k = \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mathsf{E} e^{tU}\Big|_{t=0} = \sum_{\substack{p,q \geqq 0 \\ p+q \leqq k}} \sum c_{pq}^{(k)}\, a_{pq}.$$

It remains to find the coefficent $c_{pq}^{(k)}$. From (3.2) it follows

$$\frac{\mathrm{d}}{\mathrm{d}t} I_p(t)\, I_q(t) = \tfrac{1}{2}(I_{p-1}(t)\, I_q(t) + I_{p+1}(t)\, I_q(t) + I_p(t)\, I_{q-1}(t) + I_p(t)\, I_{q+1}(t)). \tag{3.3}$$

Thus the term $I_p(t)\, I_q(t)$ "produces" the four terms on the right-hand side. Proceeding in this way we see that differentiation is analogous (in the sense of (3.3)) to producing all the paths of the symmetric random walk on $Z \times Z$ starting at the point (p, q). The probability of reaching the point $(0, 0)$ after k steps from (p, q) is of course the same as that of reaching (p, q) when starting from $(0, 0)$. Let us denote this probability by $P_k(p, q)$ and assume for a while that the starting point is $(0,0)$. Then

$$2^{-k} c_{pq}^{(k)} = P_k(p, q)\,.$$

For $P_k(p, q)$ we have

$$(3.4) \qquad P_0(0, 0) = 1\,,$$

$$P_k(p, q) = \tfrac{1}{4}[P_{k-1}(p - 1, q) + P_{k-1}(p + 1, q) + P_{k-1}(p, q - 1) + P_{k-1}(p, q + 1)]\,.$$

(3.4) defines $P_k(p, q)$ uniquely and one easily checks that it is fulfilled if $P_k(p, q)$ is given by

$$(3.5) \qquad P_k(p, q) = 4^{-k} \binom{k}{\frac{k - p - q}{2}} \binom{k}{\frac{k + p - q}{2}}$$

for both $p, k - q$ odd or both $p, k - q$ even, $P_k(p, q) = 0$ otherwise.

Let us write the first two moments:

$$\mathsf{E}U = \alpha_1(\varkappa)\,(\alpha_1(\varkappa R_1) - \alpha_1(\varkappa R_2)\cos\delta)\,,$$

$$\mathsf{E}U^2 = 1 - (\alpha_1(\varkappa R_1)\,\alpha_1(\varkappa R_2)\,(\alpha_2(\varkappa) + 1)\cos\delta) + \tfrac{1}{2}\,\alpha_2(\varkappa)\,(\alpha_2(\varkappa R_1) + \alpha_2(\varkappa R_2)\cos 2\delta)\,.$$

For practical purposes, the approximate formulas for $\alpha_1(.)$ and $\alpha_2(.)$ may be used as given in [2] 3.4.9d.

References

[1] *H. Bateman, A. Erdélyi:* Higher transcendental functions, vol. 2, McGraw-Hill 1953.

[2] *K. V. Mardia:* Statistics of directional data. Academic Press, London and New York 1972.

[3] *J. E. Morris, P. J. Laycock:* Discriminant analysis of directional data, Biometrika 61 (1974), 335—342.

Souhrn

KLASIFIKACE DO DVOU TŘÍD S VON MISESOVÝM ROZLOŽENÍM A NEZNÁMÝMI STŘEDNÍMI SMĚRY VE TŘÍDÁCH

KRYŠTOF EBEN

Článek se zabývá klasifikačním pravidlem založeným na věrohodnostním poměru a pravidlem založeným na výběrových středních směrech (plug-in rule). Pro diskriminační funkci tohoto pravidla je nalezena momentová vytvořující funkce a jsou zjištěny momenty.

Parametr koncentrace se považuje za známý a totožný v obou třídách.

Author's address: RNDr. *Kryštof Eben,* Matematický ústav ČSAV, Žitná 25, 115 67 Praha 1.