APPLICATIONES MATHEMATICAE
31,2 (2004), pp. 155–160

Stanisław Jaworski and Wojciech Zieliński (Warszawa)

A PROCEDURE FOR ε-COMPARISON OF MEANS OF TWO NORMAL DISTRIBUTIONS

Abstract. For two normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ the problem is to decide whether $|\mu_1 - \mu_2| \le \varepsilon$ for a given ε. Two decision rules are given: maximin and bayesian for σ^2 known and unknown.

1. Introduction and notation. Let X_1, X_2 be random variables with densities $f(x, \theta_1)$ and $f(x, \theta_2)$ respectively. The problem is to take one of the following decisions:

$$d_1 : |\theta_1 - \theta_2| \le \varepsilon, \qquad d_2 : |\theta_1 - \theta_2| > \varepsilon,$$

for a given $\varepsilon > 0$. This problem, in a hypothesis-testing setup, was first considered by Hodges and Lehmann [1]. Since then many authors have considered similar problems under different approaches: with ε tending to zero or sample size increasing to infinity (see for example Dette and Munk [2]). In what follows we are interested in decision-theoretic approach with given ε and finite sample size.

Consider the following procedure:

$$|T| \le t_k \;\Rightarrow\; d_1, \qquad |T| > t_k \;\Rightarrow\; d_2,$$

where T is a statistic with density $g(t, \theta)$ and $\theta = \theta_1 - \theta_2$.

Let $R(\theta, t)$ denote the probability of taking the correct decision, i.e.

$$R(\theta, t) = P_\theta\{|T| \le t\} I_\varepsilon(\theta) + P_\theta\{|T| > t\}(1 - I_\varepsilon(\theta)),$$

where $I_\varepsilon(\theta) = 0$ for $\theta \notin [-\varepsilon, \varepsilon]$ and $= 1$ for $\theta \in [-\varepsilon, \varepsilon]$. It is obvious that there does not exist a rule t_k maximizing $R(\theta, t_k)$ uniformly in θ. Hence we will find a maximin rule as well as a bayesian one.

2000 *Mathematics Subject Classification*: 62F15, 62J15.
Key words and phrases: normal distribution, ε-comparison, maximin, Bayes.

The maximin procedure t_k is a solution of the equation

$$\min_\theta R(\theta, t_k) = \max_t \min_\theta R(\theta, t).$$

The bayesian procedure t_k is a solution of the equation

$$\int R(\theta, t_k) g_a(\theta)\, d\theta = \max_t \int R(\theta, t) g_a(\theta)\, d\theta,$$

where $g_a(\theta)$ denotes an *a priori* density of θ.

2. Decision rules

LEMMA 1. *If $g(t, -\theta) = g(-t, \theta)$ and for every fixed $t > 0$ the mapping $\theta \mapsto P_\theta\{|T| \le t\}$ is a decreasing function for $\theta > 0$, then the maximin rule t_k is a solution of*

$$F_{-\varepsilon}(t_k) + F_\varepsilon(t_k) = 3/2,$$

where F_θ denotes the cdf of the statistic T.

Proof. Note that if $g(t, -\theta) = g(-t, \theta)$, then $F_{-\theta}(t) = 1 - F_\theta(-t)$. Hence $P_\theta\{|T| \le t\} = P_{-\theta}\{|T| \le t\}$. The properties of the mapping $\theta \mapsto P_\theta\{|T| \le t\}$ imply that for a given t we have

$$\forall\theta \quad R(\theta, t) \ge \min\{R(\varepsilon, t), \lim_{\vartheta \to \varepsilon^+} R(\vartheta, t)\},$$

which is seen from the following form of the probability $R(\theta, t)$:

$$R(\theta, t) = P_\theta\{|T| \le t\} I_\varepsilon(\theta) + (1 - P_\theta\{|T| \le t\})(1 - I_\varepsilon(\theta)).$$

Because for $t \ge 0$ the function $t \mapsto P_\varepsilon\{|T| \le t\}$ increases, the following equation for t_k is obtained:

$$P_\varepsilon\{|T| \le t_k\} = 1 - P_\varepsilon\{|T| \le t_k\}$$

The assertion follows by putting $P_\varepsilon\{|T| \le t_k\} = F_\varepsilon(t_k) - F_\varepsilon(-t_k) = F_\varepsilon(t_k) - 1 + F_{-\varepsilon}(t_k)$. ∎

Note that the maximin rule t_k always exists and $\min_\theta R(\theta, t_k) = 1/2$.

LEMMA 2. *If the density function $g(t, \theta)$ satisfies the following conditions*:

(i) $\forall \theta, t,\ g(t, -\theta) = g(-t, \theta)$,
(ii) *the mapping $\theta \mapsto P_\theta\{|T| \le t\}$ is a decreasing function for $\theta > 0$,*
(iii) *the mapping $t \mapsto g(t, \theta)$ is a continuous function for any fixed θ,*

then the bayesian rule is such that $t_k = 0$ or $t_k = \infty$ or t_k is a solution of the equation

$$\int_{\mathbb{R}} (I_\varepsilon(\theta) - 1/2) g(t_k, \theta) g_a(\theta)\, d\theta = 0$$

or equivalently

$$G(\varepsilon, t_k) + G(\varepsilon, -t_k) = 3/2,$$

where $G(\theta, t)$ is the a posteriori cdf of θ.

Proof. The bayesian risk of rule t equals

$$\begin{aligned}\int R(\theta,t)g_a(\theta)\,d\theta &= \int\limits_{(-\varepsilon,\varepsilon)} P_\theta\{|T|\le t\}g_a(\theta)\,d\theta + \int\limits_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} P_\theta\{|T|>t\}g_a(\theta)\,d\theta \\ &= \int\limits_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} g_a(\theta)\,d\theta + \int\limits_{\mathbb{R}} (2I_\varepsilon(\theta)-1)P_\theta\{|T|\le t\}g_a(\theta)\,d\theta\end{aligned}$$

Maximization of $\int R(\theta,t)g_a(\theta)\,d\theta$ is equivalent to maximization of

$$G(t) = \int\limits_{\mathbb{R}} (2I_\varepsilon(\theta)-1)P_\theta\{|T|\le t\}g_a(\theta)\,d\theta = \int\limits_{-t}^{t}\int\limits_{\mathbb{R}}(2I_\varepsilon(\theta)-1)g(t,\theta)g_a(\theta)\,d\theta\,dt.$$

Let $H(t) = \int_{\mathbb{R}}(2I_\varepsilon(\theta)-1)g(t,\theta)g_a(\theta)\,d\theta$. Because $I_\varepsilon(\theta) = I_\varepsilon(-\theta)$, we have $H(-t) = H(t)$. Hence

$$\frac{d}{dt}G(t) = H(t) + H(-t) = 2H(t).$$

To end the proof note that if $t_k \notin \{0,\infty\}$ then a necessary condition for a maximum at $t = t_k$ is $H(t_k) = 0$. ■

3. Illustration

Example 1. Let $X_1 \sim N(\mu_1,\sigma^2)$ and $X_2 \sim N(\mu_2,\sigma^2)$, with known σ^2. Let $\mu = \mu_1 - \mu_2$. We are going to find a procedure for taking one of the following decisions:

$$d_1 : |\mu| \le \varepsilon, \qquad d_2 : |\mu| > \varepsilon,$$

for given $\varepsilon > 0$. For samples $X_{11},\dots,X_{1n_1}$ and $X_{21},\dots,X_{2n_2}$ let

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{1/n_1 + 1/n_2}}.$$

The distribution of T is $N(\mu,1)$. Let $\varphi(\cdot)$ denote the density function and $\Phi(\cdot)$ the cdf of the $N(0,1)$ distribution.

Maximin rule. It is easy to see that φ and Φ satisfy the assumptions of Lemma 1. Hence, the maximin rule t_k is a solution of

$$\Phi(t_k - \varepsilon) + \Phi(t_k + \varepsilon) = 3/2.$$

Examples of the t_k values for different ε's are given in Table I.

Table I. Values of maximin rule t_k

ε	t_k	ε	t_k	ε	t_k
0.0	0.68696	1.0	1.06607	2.0	2.00089
0.1	0.69048	1.1	1.14592	2.1	2.10056
0.2	0.70113	1.2	1.23120	2.2	2.20035
0.3	0.71905	1.3	1.32078	2.3	2.30022
0.4	0.74450	1.4	1.41362	2.4	2.40014
0.5	0.77777	1.5	1.50880	2.5	2.50009
0.6	0.81916	1.6	1.60563	2.6	2.60006
0.7	0.86883	1.7	1.70358	2.7	2.70004
0.8	0.92678	1.8	1.80226	2.8	2.80002
0.9	0.99273	1.9	1.90142	2.9	2.90001

Bayesian rule. Let $\mu = \mu_1 - \mu_2 \sim N(0, \tau^2)$. The df φ as well as the df of the *a priori* distribution $(N(0, \tau^2))$ satisfy the assumptions of Lemma 2. Hence the bayesian rule t_k is a solution of

$$\Phi\left(\frac{t_k}{\sqrt{a}} + \sqrt{a}\,\varepsilon\right) + \Phi\left(\frac{-t_k}{\sqrt{a}} + \sqrt{a}\,\varepsilon\right) = \frac{3}{2},$$

where $a = 1/\tau^2 + 1$. Note that the solution exists if $2\Phi(\sqrt{a}\,\varepsilon) \geq 3/2$, i.e.

$$1/\tau^2 + 1 \geq \Phi^{-1}(0.75)/\varepsilon^2.$$

Let

$$\varepsilon_{\min} = \sqrt{\Phi^{-1}(0.75)\frac{\tau^2}{1+\tau^2}}$$

For $\varepsilon < \varepsilon_{\min}$ the decision d_1 is never taken.

Examples of the t_k values for different ε's are given in Table II.

Table II. Values of bayesian rule t_k

$\tau = 0.25$		$\tau = 1.00$		$\tau = 4.00$	
ε	t_k	ε	t_k	ε	t_k
0.163588	0.00000	0.476937	0.00000	0.654352	0.00000
0.20	2.66649	0.50	0.43628	0.70	0.38067
0.30	5.02746	0.60	0.98110	0.80	0.66662
0.40	6.79495	0.70	1.30029	0.90	0.85394
0.50	8.49981	0.80	1.55442	1.00	1.00441
0.60	10.20000	0.90	1.77986	1.10	1.13599
0.70	11.90000	1.00	1.99155	1.20	1.25686
0.80	13.60001	1.10	2.19667	1.30	1.37147
0.90	15.30000	1.20	2.39877	1.40	1.48238
1.00	17.00000	1.30	2.59958	1.50	1.59116
1.10	18.70001	1.40	2.79987	1.60	1.69874
1.20	20.40001	1.50	2.99996	1.70	1.80566
1.30	22.10001	1.60	3.19999	1.80	1.91223
1.40	23.80000	1.70	3.40000	1.90	2.01863
1.50	25.49999	1.80	3.60000	2.00	2.12495

EXAMPLE 2. Let $X_1 \sim N(\mu_1, \sigma^2)$ and $X_2 \sim N(\mu_2, \sigma^2)$, with unknown σ^2. Let $\mu = \mu_1 - \mu_2$. We are going to find a procedure to take one of the following decisions:

$$d_1 : |\mu| \le \varepsilon\sigma, \qquad d_2 : |\mu| > \varepsilon\sigma,$$

or equivalently

$$d_1 : |\delta| \le \varepsilon, \qquad d_2 : |\delta| > \varepsilon,$$

where $\delta = (\mu_1 - \mu_2)/\sigma$. Let

$$T = \frac{\overline{X}_1 - \overline{X}_2}{S_r},$$

where

$$S_r^2 = \frac{\sum(X_{1i} - \overline{X}_1)^2 + \sum(X_{2i} - \overline{X}_2)^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right).$$

The statistic T has noncentral t-distribution with noncentrality parameter δ and $n_1 + n_2 - 2$ degrees of freedom.

Maximin rule. It is easy to check that the density function and cdf of the noncentral t distribution satisfy the assumptions of Lemma 1. Hence the maximin rule t_k is a solution of the equation

$$T_{\nu,-\varepsilon}(t_k) + T_{\nu,\varepsilon}(t_k) = 3/2,$$

where $T_{\nu,\delta}(\cdot)$ denotes the cdf of the t distribution with $\nu = n_1 + n_2 - 2$ degrees of freedom and noncentrality parameter δ.

Examples of the t_k values for different ε's are given in Table III.

Table III. Values of maximin rule t_k

ε	t_k	ε	t_k	ε	t_k
0.0	0.68695	1.0	1.06769	2.0	2.02613
0.1	0.69040	1.1	1.14981	2.1	2.12741
0.2	0.70079	1.2	1.23774	2.2	2.22876
0.3	0.71833	1.3	1.33017	2.3	2.33013
0.4	0.74334	1.4	1.42584	2.4	2.43152
0.5	0.77620	1.5	1.52371	2.5	2.53294
0.6	0.81734	1.6	1.62302	2.6	2.63436
0.7	0.86708	1.7	1.72321	2.7	2.73580
0.8	0.92556	1.8	1.82392	2.8	2.83724
0.9	0.99262	1.9	1.92494	2.9	2.93870

Bayesian rule. Let $\delta \sim N(0, \tau^2)$. The density function of the noncentral t distribution as well as the df of the *a priori* distribution ($N(0, \tau^2)$) satisfy the assumptions of Lemma 2. Hence the bayesian rule t_k is a solution of

$$T_{\nu+1,-\sqrt{a}\varepsilon}\left(\frac{\sqrt{\nu+1}}{h(t_k)}\right) + T_{\nu+1,-\sqrt{a}\varepsilon}\left(-\frac{\sqrt{\nu+1}}{h(t_k)}\right) = \frac{3}{2},$$

where $h(t_k)^2 = a(\nu/t_k^2 + 1) - 1$. Let

$$\varepsilon_{\min} = \Phi^{-1}(0.75)/\sqrt{a},$$

$$\varepsilon_{\max} = \max\{\varepsilon : T_{\nu+1,-\varepsilon\sqrt{a}}(\tau\sqrt{\nu+1}) + T_{\nu+1,-\varepsilon\sqrt{a}}(-\tau\sqrt{\nu+1}) \le 1.5\}.$$

For $\varepsilon < \varepsilon_{\min}$ the decision d_1 is never taken. For $\varepsilon > \varepsilon_{\max}$ the decision d_2 is never taken.

Examples of the t_k values for different ε's are listed in Table IV.

Table IV. Values of bayesian rule t_k

$\tau = 0.25$		$\tau = 1$		$\tau = 4$	
ε	t_k	ε	t_k	ε	t_k
0.163588	0.00000	0.476937	0.00000	0.654352	0.00000
0.1810	1.99301	0.7690	1.49637	2.3800	2.52191
0.2030	3.35621	1.3070	2.82801	5.7400	6.38610
0.2250	4.88682	1.8450	4.46261	9.1000	11.20500
0.2360	5.88206	2.1140	5.56649	10.7800	14.39483
0.2470	7.20943	2.3830	7.06078	12.4600	18.66551
0.2580	9.23189	2.6520	9.37839	14.1400	25.23805
0.2690	13.22997	2.9210	14.18499	15.8200	38.86985
0.2800	36.20679	3.1900	91.05546	17.5000	601.61613
0.281886	∞	3.19767	∞	17.5082	∞

Acknowledgements. The authors would like to thank Professor Teresa Ledwina for a helpful discussion.

References

[1] J. L. Hodges and E. L. Lehmann, *Testing the approximate validity of statistical hypotheses*, Ann. Math. Statist. 16 (1954), 261–268.

[2] H. Dette and A. Munk, *Validation of linear models*, Ann. Statist. 26 (1998), 778–800.

Department of Econometrics and Computer Science
Warsaw Agricultural University
Nowoursynowska 166
02-787 Warszawa, Poland
E-mail: stanislaw.jaworski@ekonometria.info
wojtek.zielinski@statystyka.info

Received on 16.4.2003;
revised version on 20.11.2003 (1681)