Journal of the Korean
Statistical Society
Vol. 22, No. 1, 1993

Detection of the Normal Population with the Largest Absolute Value of Mean†

Woo-Chul Kim[1] and Gyu-Jin Jeong[2]

ABSTRACT

Among k independent normal populations with unknown means and a common unknown variance, the problem of detecting the population with the largest absolute value of mean is considered. This problem is formulated in a manner close to the framework of testing hypotheses, and the maximum error probability and the minimum power are considered. The power charts necessary to determine the sample size are provided. The problem of detecting the population with the smallest absolute value of mean is also considered.

KEYWORDS: Detection of the normal population, lagest absolute value of mean, maximum error probability, minimum power, power charts.

---

[1] Department of Computer Science and Statistics, Seoul National University, Seoul, 151-742, Korea.
[2] Department of Applied Statistics, Hannam University, Daejeon 300-791, Korea.
† This research was supported by SNU Development Foundation 1987-1988.

1. INTRODUCTION

We observe independent X_{ij} $(i = 1, \cdots, k, j = 1, \cdots, n)$ from each of k normal populations $N(\mu_i, \sigma^2)$ with common variance, and let $\overline{X}_i$ $(i = 1, \cdots, k)$ and $\hat{\sigma}^2$ denote the sample means and the pooled sample variance, respectively. Suppose we are interested in detecting the *best* population, that corresponds to the largest $\theta_i = |\mu_i|/\sigma$, the so-called signal-to-noise ratio.

One would naturally order the values of $Y_i = |\overline{X}_i|/\hat{\sigma}$ $(i = 1, \cdots, k)$, and make an inference

$$\theta_{(k)} > \max_{i \neq (k)} \theta_i \quad if \quad Y_{(k)} > \max_{i \neq (k)} Y_i + c\sqrt{2/n}, \tag{1.1}$$

where $Y_{(1)} \leq \cdots \leq Y_{(k)}$ denote the ordered Y_i's $(i = 1, \cdots, k)$. Then the probability of a false inference is

$$P_{\underline{\mu},\sigma}\left(\theta_{(k)} \leq \max_{i \neq (k)} \theta_i,\ Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n} \right), \tag{1.2}$$

and the probability of a correct inference is

$$P_{\underline{\mu},\sigma}\left(\theta_{(k)} > \max_{i \neq (k)} \theta_i,\ Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n} \right). \tag{1.3}$$

In the sequel, the probabilities in (1.2) and (1.3) will be simply called an error probability and a power, respectively.

The goal of this paper is to find the maximum error probability in (1.2) and to determine the sample size to guarantee the power in (1.3). The formulation in this paper is that of Gutmann and Maymin (1987), who adopted it in location parameters problem. Similar formulations have been adopted in Gutmann (1985), Bofinger (1988), Kim (1988), and Jeon, Kim and Jeong (1988) for various comparison problems.

Finally it should be mentioned that the selection problem in terms of $|\mu_i|/\sigma$ was studied by Rizvi (1971) under the frameworks of Bechhofer (1954) and Gupta (1956). Comparison with Rizvi's (1971) will be given in section 3.

2. MAIN RESULTS

We first find the maximum error probability of the procedure (1.1) using the following two lemmas, whose proofs are deferred to section 4.

LEMMA 1. The error probability in (1.2) is maximized when at least two θ_i's are tied as the largest.

LEMMA 2. When exactly r $(2 \leq r \leq k)$ θ_i's are tied as the largest, the maximum error probability is given by

$$\int_0^{\infty} \int_{-\infty}^{\infty} r\Phi^{r-1}(y - \sqrt{2}cu)d\Phi(y)dQ_\nu(u), \tag{2.1}$$

where Φ and Q_ν denote the cdf's of $N(0,1)$ and $\sqrt{\chi^2(\nu)/\nu}$ with $\nu = k(n-1)$.

It follows from the above lemmas that the maximum error probability is found by taking the maximum of (2.1) with respect to $r = 2, \cdots, k$. It can be, however, easily observed that (2.1) is non-increasing in r. Thus we have the next result.

THEOREM 1. For the procedure (1.1), the maximum error probability is given by

$$\begin{aligned} &\max_{\underline{\mu},\sigma} P_{\underline{\mu},\sigma}\ (\theta_{(k)} \leq \max_{i \neq (k)} \theta_i,\ Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n}) \\ &= 2\int_0^{\infty} \int_{-\infty}^{\infty} \Phi(y - \sqrt{2}cu)d\Phi(y)dQ_\nu(u). \end{aligned} \tag{2.2}$$

It turns out from Theorem 1 that the choice of $c = t_{\alpha/2}(\nu)$, the upper $\alpha/2$ quantile of t-distribution with ν degrees of freedom, controls the error probability below α. Next, let us consider the power in (1.3) of the procedure (1.1). The region of interest, where the power is controlled, is taken as

$$\Omega(\delta) = \{\ (\underline{\mu}, \sigma)\ :\ (\theta_{[k]} - \theta_{[k-1]}) \geq \delta\ \} \tag{2.3}$$

where $\delta > 0$ is specified by the experimenter and $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$ are the ordered θ_i's $(i = 1, \cdots, k)$. The region $\Omega(\delta)$ given by (2.3) is the so-called preference zone introduced by Bechhofer (1954) in his indifference zone approach, and represents the case where the detection of the best is really desired.

For any $(\underline{\mu}, \sigma) \in \Omega(\delta)$, the power of the procedure (1.1) is given by

$$P_{\underline{\mu},\sigma}(\ \theta_{(k)} > \max_{i \neq (k)} \theta_i,\ Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n}\ )$$
$$= P_{\underline{\mu},\sigma}(\ Y_{[k]} = Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n}\ )\ , \tag{2.4}$$

where $Y_{[k]}$ is the Y_i associated with $\theta_{[k]} = \max_{1 \leq i \leq k} \theta_i$.

By methods similar to Lemmas 1 and 2, it can be shown that the power in (2.4) is minimized when $\theta = \theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \delta$ and the common value θ approaches $+\infty$. Thus we state the next result without proof.

THEOREM 2. For the procedure (1.1), the minimum power over $\Omega(\delta)$ is given by

$$\min_{\theta_{[k]} - \theta_{[k-1]} \geq \delta} P_{\underline{\mu},\sigma}\ (\theta_{(k)} > \max_{i \neq (k)} \theta_i,\ Y_{(k)} > Y_{(k-1)} + c\sqrt{2/n})$$
$$= \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(y + \sqrt{n}\delta - \sqrt{2}cu) d\Phi(y) dQ_\nu(u). \tag{2.5}$$

For given k, α, and δ, the minimum power (2.5) can be computed as a function of sample size n. This has been done and the power charts are given in Figure 1 for $\alpha = 0.05, \delta = 0.5,\ 0.7,\ 1.0$ and $k = 3,\ 4,\ 5,\ 6$.

Figure 1
Minimum power vs sample size for alpha=0.05

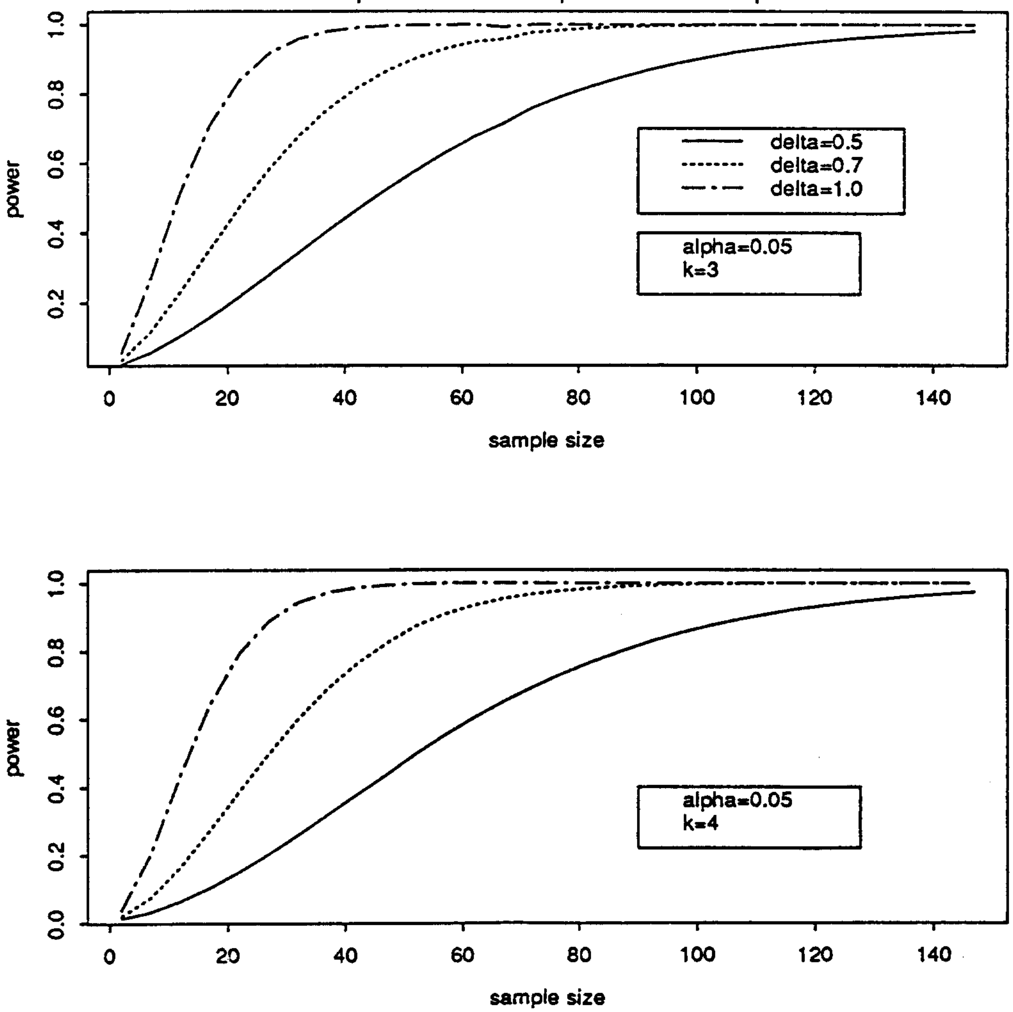
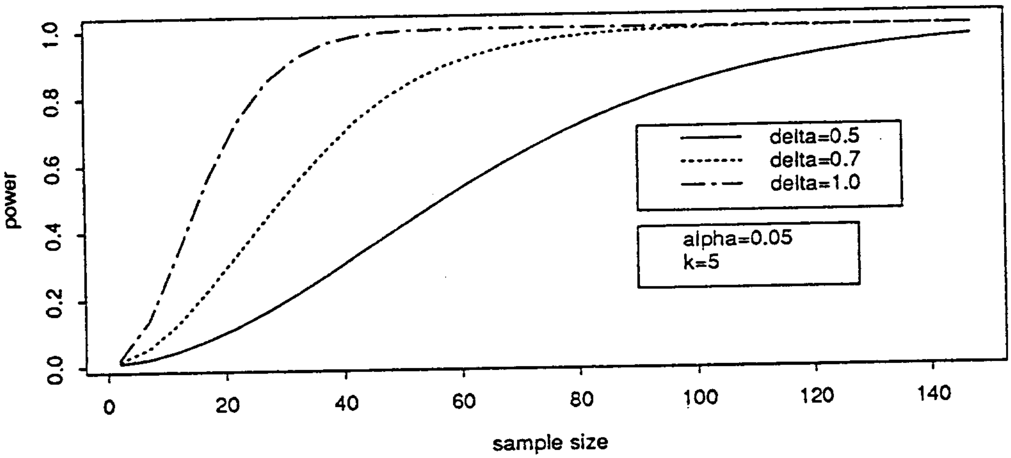
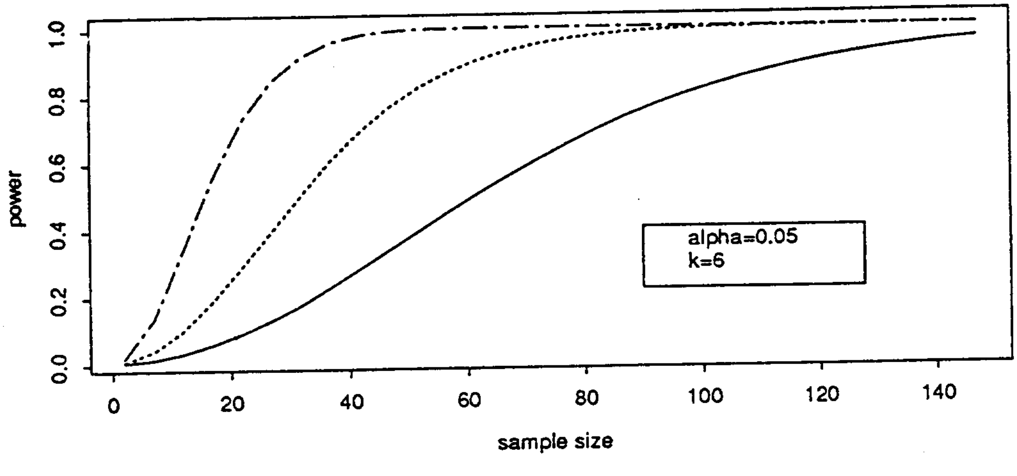


Figure 1 (continued)





To construct Figure 1, numerical evaluation of the double integral in (2.5) was done by using the 64 points Gauss-Hermite quadrature for the inner integral and 32 points Gauss-Laguerre quadrature for the outer integral, both in the IBM Scientific Subroutine Package. The values of normal cdf in the inner integral were obtained via IMSL's subroutine MDNOR.

3. CONCLUDING REMARKS

Rizvi (1971) considered the selection problem in terms of $|\mu_i|/\sigma$ under Gupta's (1956) subset selection framework. His procedure is to

$$\text{select population } j \text{ if } Y_j > \max_{i \neq j} Y_i + c^G\sqrt{2/n},$$

where c^G is the Gupta's constant for normal means problem. He showed that his procedure guarantees the probability of *including the best* in the selected subset to be at least $1 - \alpha$.

Due to different formulations, it is difficult to compare the procedure (1.1) with Rizvi's. But a comparison of the critical values $t_{\alpha/2}(\nu)$ with c^G (see, for example, Gupta, Panchapakesan and Sohn (1985)) shows that $t_{\alpha/2}(\nu) < c^G$ for $k > 3$ and for $\alpha = 0.05,\ 0.01$, regardless of the degrees of freedom ν. Thus it can be said that, in detecting *only* the best, the procedure (1.1) makes the useful inference $\theta_{(k)} > \max_{i \neq (k)} \theta_i$ more often than Rizvi's for $k > 3$.

Finally, it should be remarked that the same formulation and similar procedure can be devised to detect the normal population with the smallest $|\mu_i|/\sigma$. Such a problem is of interest when we wish to detect the population with the least fraction defective in the presence of both upper and lower specifications.

With similar techniques, it can be shown that, for the procedure

$$\text{“ } \theta_{(1)} < \min_{i \neq (1)} \theta_i \text{ if } Y_{(1)} < \min_{i \neq (1)} Y_i - c\sqrt{2/n} \text{ ”},$$

the maximum error probability and the minimum power when $\theta_{[2]} - \theta_{[1]} \geq \delta$ are given by (2.2) and (2.5), respectively. Thus the design constant $t_{\alpha/2}(\nu)$ and the power charts Figure 1 can be also used in such a problem.

4. PROOFS OF LEMMAS

First note that in proving Lemmas 1 and 2 we may assume $\hat{\sigma}/\sqrt{n} = \sigma/\sqrt{n} = 1$ by considering the conditional error probability given the sample variance $\hat{\sigma}^2$. Thus we are assuming $\hat{\sigma}/\sqrt{n} = \sigma/\sqrt{n} = 1$ in the sequel, and let $D = \sqrt{2}c$ for notational convenience.

LEMMA 1. By symmetry, we may assume that $\theta_1 \le \cdots \le \theta_k$. Note that the event $\{\theta_{(k)} \le \max_{i \neq (k)} \theta_i\}$ always occurs when $\theta_k = \theta_{k-1}$. Thus the maximum error probability when $\theta_k = \theta_{k-1}$ is given by

$$\sup_{\theta \in \Omega_0} P_{\underline{\theta}}(Y_{(k)} > Y_{(k-1)} + D), \tag{4.1}$$

where $\Omega_0 = \{\theta : \theta_1 \le \cdots \le \theta_{k-1} = \theta_k\}$.

When $\theta_k > \theta_{k-1}$, the error probability is given by

$$P_{\underline{\theta}}(\max_{1 \le j \le k-1} Y_j = Y_{(k)} > Y_{(k-1)} + D)$$

which is clearly non-increasing in θ_k. Thus it is maximized as θ_k decreases to θ_{k-1}. Therefore the maximum error probability when $\theta_k > \theta_{k-1}$ is given by

$$\sup_{\underline{\theta} \in \Omega_0} P_{\underline{\theta}}(\max_{1 \le j \le k-1} Y_j = Y_{(k)} > Y_{(k-1)} + D)$$

which is clearly not larger than (4.1). This completes the proof.

To prove Lemma 2, note that for any $\theta \in \Omega_0$ the error probability is given by

$$P_{\underline{\theta}}(Y_{(k)} > Y_{(k-1)} + D) = \sum_{j=1}^{k} \int_0^{\infty} \prod_{\substack{i=1 \\ i \neq j}}^{k} F(y - D, \theta_i)\, f(y, \theta_j)\, dy, \tag{4.2}$$

where $F(y, \theta_i)$ and $f(y, \theta_i)$ denote the cdf and the pdf of $Y_i (i = 1, \cdots, k)$, respectively. Since we assumed $\hat{\sigma}/\sqrt{n} = \sigma/\sqrt{n} = 1$, the cdf and the pdf of Y_i are given by

$$F(y, \theta_i) = \Phi(y - \theta_i) - \Phi(-y - \theta_i) \text{ and } f(y, \theta_i) = \phi(y - \theta_i) + \phi(y + \theta_i).$$

Then, by observing that, for all $\theta \ge \theta_1 \ge 0$, $\{\phi(y+\theta) - \phi(y-\theta)\}/\{\phi(y+\theta_1) + \phi(y - \theta_1)\}$ is non-increasing in y, which can be easily shown by differentiation, and by simple differentiation, we have

$$\frac{\partial}{\partial \theta} F(y - D, \theta)\, f(y, \theta_1) - \frac{\partial}{\partial \theta} F(y, \theta)\, f(y - D, \theta_1) \ge 0 \tag{4.3}$$

for any $D > 0$.

LEMMA 2. By symmetry we may assume $\theta_1 \le \cdots \le \theta_{k-r} < \theta_{k-r+1} = \cdots = \theta_k = \theta$. Then the error probability (4.2) is given by

$$\begin{aligned} & P_{\underline{\theta}}(\, Y_{(k)} > Y_{(k-1)} + D \,) \\ = \; & r \int_D^\infty F^{r-1}(y-D,\theta) \prod_{i=1}^{k-r} F(y-D,\theta_i)\, f(y,\theta)\, dy \\ & + \sum_{j=1}^{k-r} \int_D^\infty F^r(y-D,\theta) \prod_{\substack{i=1 \\ i\ne j}}^{k-r} F(y-D,\theta_i)\, f(y,\theta_j)\, dy. \end{aligned}$$

By differentiating this expression w.r.t θ, we have

$$\begin{aligned} & \frac{\partial}{\partial\theta} P_{\underline{\theta}}(\, Y_{(k)} > Y_{(k-1)} + D \,) \\ = \; & r(r-1) \int_D^\infty F^{r-2}(y-D,\theta) \prod_{i=1}^{k-r} F(y-D,\theta_i) \\ & \qquad \left[\frac{\partial}{\partial\theta} F(y-D,\theta)\, f(y,\theta) - \frac{\partial}{\partial\theta} F(y,\theta)\, f(y-D,\theta)\right] dy \\ & + \sum_{j=1}^{k-r} r \int_D^\infty F^{r-1}(y-D,\theta) \prod_{\substack{i=1 \\ i\ne j}}^{k-r} F(y-D,\theta_i) \\ & \qquad \left[\frac{\partial}{\partial\theta} F(y-D,\theta)\, f(y,\theta_j) - \frac{\partial}{\partial\theta} F(y,\theta)\, f(y-D,\theta_j)\right] dy. \end{aligned}$$

Since the expressions in brackets are non-negative by (4.3), the error probability is maximized as θ approaches $+\infty$. Thus, it suffices to find the maximum error probability when θ approaches $+\infty$. When the largest r parameters are tied, the error probability can be computed in the following way :

$$\begin{aligned} & P_{\underline{\theta}}(\, Y_{(k)} > Y_{(k-1)} + D \,) \\ = \; & r \int_D^\infty \{\Phi(y-D-\theta) - \Phi(-y+D-\theta)\}^{r-1} \\ & \prod_{i=1}^{k-r} \{\Phi(y-D-\theta_i) - \Phi(-y+D-\theta_i)\}\, \{\phi(y-\theta) + \phi(y+\theta)\}\, dy \\ & + \sum_{j=1}^{k-r} \int_D^\infty \{\Phi(y-D-\theta) - \Phi(-y+D-\theta)\}^r \end{aligned}$$

$$\prod_{\substack{i=1\\i\neq j}}^{k-r} \{\Phi(y-D-\theta_i)-\Phi(-y+D-\theta_i)\}\{\phi(y-\theta_j)+\phi(y+\theta_j)\}\,dy$$

$$= r\int_{D-\theta}^{\infty}\{\Phi(y-D)-\Phi(-y+D-2\theta)\}^{r-1}$$

$$\prod_{i=1}^{k-r}\{\Phi(y+\theta-D-\theta_i)-\Phi(-y-\theta+D-\theta_i)\}\,\phi(y)\,dy$$

$$+r\int_{D+\theta}^{\infty}\{\Phi(y-D-2\theta)-\Phi(-y+D)\}^{r-1}$$

$$\prod_{i=1}^{k-r}\{\Phi(y-\theta-D-\theta_i)-\Phi(-y+\theta+D-\theta_i)\}\,\phi(y)\,dy$$

$$+\sum_{i=1}^{k-r}\int_{D}^{\infty}\{\Phi(y-D-\theta)-\Phi(-y+D-\theta)\}^{r}$$

$$\prod_{\substack{i=1\\i\neq j}}^{k-r}\{\Phi(y-D-\theta_i)-\Phi(-y+D-\theta_i)\}\{\phi(y-\theta_j)+\phi(y+\theta_j)\}\,dy.$$

By sending θ to $+\infty$ and by taking the expectation with respect to $\hat{\sigma}/\sigma$, we can easily observe that the maximum error probability is given by (2.1).

REFERENCES

(1) Bechhofer, R.E. (1954). A single sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, 25, 16-39.

(2) Bofinger, E. (1988). The least significant spacing for " one versus the rest " normal populations. *Communications in Statistics, -Theory and Methods,* A17, 1697-16.

(3) Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeo. Ser No. 150), Institute of Statistics, University of North Carolina, Chapel Hill.

(4) Gupta, S.S., Panchapakesan, S., and Sohn, J.K. (1985). On the distribution of the Studentized maximum of equally correlated normal random variables. *Communications in Statistics,* B14, 103-135.

(5) Gutmann, S. (1985). Is the selected population unsurpassed ? (Three or four normal populations), Unpublished manuscript, Northesatern University, U.S.A.

(6) Gutmann, S. and Maymin, Z. (1987). Is the selected population the best? *Annals of Statistics*, 15, 456-461.

(7) Jeon, J.W., Kim, W.C., and Jeong, G.J. (1988). Confidence bounds for superiority. *Journal of The Korean Society for Quality Control*, 16, No.2, 10-16.

(8) Kim, W.C. (1988). On detecting the best treatment. *Journal of The Korean Statistical Society*, 17, 82-92.

(9) Rizvi, M.H. (1971). Some selection problems involving folded normal distribution. *Technometrics*, 13, 355-369.