Advances and Applications in Statistics
Volume 9, Number 1, 2008, Pages 127-139
This paper is available online at http://www.pphmj.com

SIMULTANEOUS NONPARAMETRIC CONFIDENCE REGIONS FOR COMPARING SUCCESSIVE TREATMENTS

YUH-ING CHEN and KUANG-LU WANG

Institute of Statistics
National Central University
Jhongli 320
Taiwan
e-mail: ychen@stat.ncu.edu.tw

Abstract

In this paper, we are concerned with multiple comparison problems in a one-way layout under the general setting with continuous population distributions. Of interest are the comparisons of successive treatments when the treatment effects are expected to be increasing or decreasing with treatment levels. Let X_i be the random variable corresponding to the ith group, $i=1,...,k$. To make the comparisons, we construct simultaneous confidence regions for probabilities $P(X_i < X_{i+1})$, $i = 1, ..., k-1$. A simulation study is conducted to investigate the coverage probability of the proposed simultaneous lower confidence bounds for the probability and the ones suggested by Lee and Spurrier [9] for comparing successive population medians. Finally, the proposed confidence region is illustrated by using a data set from a subchronic study of a chemical on liver damage conducted by the National Toxicology Program.

2000 Mathematics Subject Classification: 62.

Keywords and phrases: Mann-Whitney statistic, multiple comparison procedure, simultaneous confidence region, successive treatments.

Received January 30, 2008

1. Introduction

The effects of a substance (e.g., a toxin or a drug) are usually investigated by an experiment including several increasing or decreasing levels of the substance. It often assumes reasonably that the larger dose level produces stronger or at least equal treatment effects. However, the exact dose response relationship remains unknown. To get insight the real pattern for the dose response relationship, Lee and Spurrier [8] and Liu et al. [11] considered multiple comparisons between adjacent treatments to decide if an increasing or decreasing level leads to an additional effect when the underlying populations are normally distributed. However, the assumption of normal distributions may not be feasible in practical world. Lee and Spurrier [9] then suggested nonparametric procedures for the successive-treatment comparisons when the underlying continuous distributions are different only in location parameters.

Most of the multiple comparison procedures developed under a one-way layout involve several two-sample statistics. For example, Lee and Spurrier [8] and Liu et al. [11] employed several two-sample student's t statistics and Lee and Spurrier [9] considered several two-sample Mann-Whitney [12] statistics. However, for the general two-sample problem, it occurs quite often that the two populations may be different both in location and scale. Therefore, Sen [14] estimated the variance of the Mann-Whitney statistic and constructed a confidence interval for the probability that the variable (X_1) in one group is smaller than that (X_2) in the other group, denoted by $P(X_1 < X_2)$. Halperin et al. [6] simplified the variance estimate of the Mann-Whitney statistic as a function of $P(X_1 < X_2)$ and then obtained a confidence region for the probability following the procedure in Ghosh [5] for a binomial proportion. However, the variance estimate developed in Halperin et al. [6] underestimates the true variance and, hence, the confidence interval suggested therein tends to be anti-conservative in holding its confidence level. Therefore, Mee [13] employed a variance estimate in Sen [14] to construct a binomial proportion-type confidence interval or bound for $P(X_1 < X_2)$. In fact, the

confidence bound suggested in Mee [13] maintains the confidence level better than that of the one proposed in Halperin et al. [6]. Moreover, both the binomial proportion-type lower confidence bounds are superior to the one in Sen [14] on holding the confidence level.

Therefore, in this paper, we consider to contrast adjacent treatment effects based on Mee's [13] procedure. Under the one-way layout, let $X_1, ..., X_k$ be the independent random variables distributed as continuous distribution functions $F_1(x), ..., F_k(x)$, respectively. In Section 2, for comparing adjacent treatments under such a general setting, we develop a simultaneous confidence region for probabilities $P(X_i < X_{i+1})$, $i = 1, ..., k-1$, when treatment effects are expected to have an increasing or decreasing order. In Section 3, we present and discuss the results of a simulation study for contrasting the proposed simultaneous lower confidence bounds with the ones suggested in Lee and Spurrier [9] for a variety of ordered treatment effects configurations under normal and two-parameter exponential distribution. Finally, in Section 4, we demonstrate the application of the proposed procedure by illustrating a data set from a subchronic study conducted by the National Toxicology Program investigation the effect of vinylidene fluoride on liver damage (Dietz [3]).

2. Proposed Confidence Regions

We consider the multiple comparisons between adjacent treatments by constructing simultaneous confidence regions for $p_i = P(X_i < X_{i+1}) = \int F_i dF_{i+1}, i = 1, ..., k-1$. Let $X_{i1}, ..., X_{in_i}, i = 1, ..., k$, be k independent random samples. A consistent and unbiased estimator for p_i is then given by $\hat{p}_i = \int \hat{F}_i d\hat{F}_{i+1}$, where $\hat{F}_i(x) = n_i^{-1} \sum_{j=1}^{n_i} \phi(X_{ij}, x)$ with $\phi(a, b) = 1$, if $a < b$, $= 0$, otherwise. Let $R_{i+1,j}^{(i)}$ denote the rank of $X_{i+1,j}$ among all the observations in the combined samples of the ith and $(i+1)$th samples. Set $\overline{R}_{i+1}^{(i)} = n_{i+1}^{-1} \sum_{j=1}^{n_{i+1}} R_{i+1,j}^{(i)}$. The equivalence of the Mann-Whitney statistic and Wilcoxon's [15] average rank further leads to

$$\hat{p}_i = \frac{1}{n_i}\left(\overline{R}_{i+1}^{(i)} - \frac{n_{i+1}+1}{2}\right). \tag{2.1}$$

Note that $\text{var}(\hat{p}_i)$ can be expressed as $p_i(1-p_i)/N_i$ which can be consistently estimated (Sen [14]) by

$$\widehat{\text{var}}(\hat{p}_i) = \frac{1}{n_i n_{i+1}}\left[\frac{1}{n_i(n_{i+1}-1)}\sum_{j=1}^{n_{i+1}}\left(R^{(i)}_{i+1,j} - R_{i+1,j} - \bar{R}^{(i)}_{i+1} + \frac{n_{i+1}+1}{2}\right)^2\right.$$

$$\left. + \frac{1}{n_{i+1}(n_i-1)}\sum_{j=1}^{n_i}\left(R^{(i)}_{ij} - R_{ij} - \bar{R}^{(i)}_{i} + \frac{n_i+1}{2}\right)^2\right], \quad (2.2)$$

where R_{ij} is the rank of X_{ij} among all the observations within the ith sample. Therefore, the N_i can be estimated by

$$\hat{N}_i = [\hat{p}_i(1-\hat{p}_i)]/\widehat{\text{var}}(\hat{p}_i). \quad (2.3)$$

The binomial proportion-type confidence interval for one particular p_i (Mee [13]) can then be solved by setting

$$\frac{\sqrt{\hat{N}_i}\,|\,\hat{p}_i - p_i\,|}{\sqrt{p_i(1-p_i)}} \le z_{\alpha/2},$$

where z_a is the upper ath percentile of a standard normal distribution.

We now construct a simultaneous confidence region for the p_i's. To do so, we obtain the following result according to the Central Limit Theorem and Continuity Theorem (Billingsley [1]):

Theorem. *Let* $n_T = \sum_{i=1}^{k} n_i$. *Suppose* $n_i/n_T \to \lambda_i$, *for some constant* $0 < \lambda_i < 1$, *as* $n_T \to \infty$, $i = 1, \ldots, k$. *Then as* $n_T \to \infty$

$$\left(\frac{\sqrt{\hat{N}_1}(\hat{p}_1 - p_1)}{\sqrt{p_1(1-p_1)}}, \ldots, \frac{\sqrt{\hat{N}_{k-1}}(\hat{p}_{k-1} - p_{k-1})}{\sqrt{p_{k-1}(1-p_{k-1})}}\right) \xrightarrow{d} N_{k-1}(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma} = (\rho_{ij})_{k-1\times k-1}$ *and* $\rho_{ii} = 1$ *for* $i = 1, \ldots, k-1$, $\rho_{i,i+l} = 0$ *for* $l \ge 2$ *and* $\rho_{i,i+1} = \lim_{n_T\to\infty} \text{corr}(\hat{p}_i, \hat{p}_{i+1})$.

Note that, following Chen and Wolfe [2], the covariance can be estimated by

$$\widehat{\mathrm{cov}}(\hat{p}_i, \hat{p}_{i+1}) = \frac{-1}{n_{i+1}(n_{i+1}-1)} \sum_{j=1}^{n_{i+1}} \left[\frac{1}{n_i} \left(R_{i+1,j}^{(i)} - R_{i+1,j} - \bar{R}_{i+1}^{(i)} + \frac{n_{i+1}+1}{2} \right) \right]$$

$$\left[\frac{1}{n_{i+2}} \left(R_{i+1,j}^{(i+1)} - R_{i+1,j} - \bar{R}_{i+1}^{(i+1)} + \frac{n_{i+1}+1}{2} \right) \right].$$

An estimate of $\rho_{i,i+1}$ is then given by

$$\hat{\rho}_{i,i+1} = \frac{\widehat{\mathrm{cov}}(\hat{p}_i, \hat{p}_{i+1})}{\sqrt{\widehat{\mathrm{var}}(\hat{p}_i)\,\widehat{\mathrm{var}}(\hat{p}_{i+1})}},$$

where $\widehat{\mathrm{var}}(\hat{p}_i)$ is stated in (2.2). Moreover, substituting, $\rho_{i,i+1}$ by $\hat{\rho}_{i,i+1}$, we have an estimate of $\boldsymbol{\Sigma}$, denoted by $\hat{\boldsymbol{\Sigma}}$.

Let $z_{\alpha,1}$ and $z_{\alpha,2}$ be the upper αth percentiles of the distributions of $\max(Z_1, \ldots, Z_{k+1})$ and $\max(|Z_1|, \ldots, |Z_{k-1}|)$, respectively, where $(Z_1, \ldots, Z_{k+1})$ is distributed as $N_{k-1}(\mathbf{0}, \hat{\boldsymbol{\Sigma}})$. Note that $z_{\alpha,1}$ and $z_{\alpha,2}$ can be solved by using the program in Drezner [4]. Therefore, we can employ the values of $z_{\alpha,1}$ or $z_{\alpha,2}$ to construct approximate simultaneous confidence bounds or intervals for the p_i's. The approximate level-$(1-\alpha)$ simultaneous confidence intervals for the p_i's are then obtained as

$$(1 + z_{\alpha,2}^2/\hat{N}_i)^{-1}\{\hat{p}_i + 0.5z_{\alpha,2}^2/\hat{N}_i \pm \sqrt{z_{\alpha,2}^2/\hat{N}_i}[\hat{p}_i(1-\hat{p}_i) + 0.25z_{\alpha,2}^2/\hat{N}_i]^{1/2}\},$$

where the $\hat{N}_i$ is given in (2.3). Moreover, the approximate level-$(1-\alpha)$ simultaneous lower and upper confidence bounds for the p_i's are given by

$$\hat{p}_i^L = (1 + z_{\alpha,1}^2/\hat{N}_i)^{-1}\{\hat{p}_i + 0.5z_{\alpha,1}^2/\hat{N}_i - \sqrt{z_{\alpha,1}^2/\hat{N}_i}[\hat{p}_i(1-\hat{p}_i) + 0.25z_{\alpha,1}^2/\hat{N}_i]^{1/2}\}$$

and

$$\hat{p}_i^U = (1 + z_{\alpha,1}^2/\hat{N}_i)^{-1}\{\hat{p}_i + 0.5z_{\alpha,1}^2/\hat{N}_i + \sqrt{z_{\alpha,1}^2/\hat{N}_i}[\hat{p}_i(1-\hat{p}_i) + 0.25z_{\alpha,1}^2/\hat{N}_i]^{1/2}\},$$

respectively.

Note that, to perform the one-sided inference, we need to compute $z_{\alpha,1}$ such that the probability $P(\max(Z_1, ..., Z_{k-1}) \geq z_{\alpha,1}) = \alpha$, where $(Z_1, ..., Z_{k-1})$ is distributed to $N_{k-1}(\mathbf{0}, \Sigma)$, by performing a suitable program. However, if the program is not available, according to Hunter [7] or Worsley [16], the probability can then be bounded above by

$$(k-1)P(Z_1 > z_{\alpha,1}) - \sum_{i=1}^{k-2} P(Z_i > z_{\alpha,1}, Z_{i+1} > z_{\alpha,1}).$$

Let $\Phi(.)$ be the distribution function of a standard normal random variable. Lee and Spurrier [8] modified the upper bound and suggested use of

$$(k-1)[1-\Phi(z_{\alpha,1})] - (k-2)[1-\Phi(z_{\alpha,1})]^2$$

denoted by LS, which is, in fact, a probability upper bound when the number of treatments $k \geq 5$. Due to the particular covariance structure, the joint probability $P(Z_i > z_{\alpha,1}, Z_{i+1} > z_{\alpha,1})$ can be neglected for large $z_{\alpha,1}$, but the remaining joint two-variable probability should be taken into account. Therefore, following the results in, again, Hunter [7] or Worsley [16], we obtain an upper bound for the probability,

$$(k-1)[1-\Phi(z_{\alpha,1})] - [(k-2)(k-3)/2][1-\Phi(z_{\alpha,1})]^2,$$

denoted by CW, for $k = 3, 4$ and 5. In fact, CW bound reduces to the usual Bonferroni upper bound when $k = 3$ and coincides with the LS bound when $k = 5$.

To investigate the appropriateness of the upper bounds, we computed, by using the program in Drezner [4], the critical values $z_{\alpha,1}$ for $1-\alpha = 0.90$ or 0.95, $3 \leq k \leq 7$ and equi-correlation coefficient $\rho_{i,i+1} = \rho$ ranges from –0.1 to –0.5, which occurs often in practice. The probability upper bounds, reported in Table 1, indicate that the CW bound provides a better approximation for the probability than the LS bound for all k under study, except for $k = 5$, when the involved random variables are

equi-correlated. In general, to obtain approximated critical values only based on the $\Phi(.)$, use of CW bound is suggested for $k = 3$ or 4 and LS bound is recommended for $k \geq 5$.

Table 1. Comparisons of probabilities for one-side critical values obtained CW and LS bounds for k treatment groups with equi-correlation coefficient $\rho_{i,i+1} = \rho$

$1-\alpha$	k	Method	ρ				
			–0.1	–0.2	–0.3	–0.4	–0.5
0.90	3	CW	0.9016	0.9009	0.9005	0.9002	0.9001
		LS	0.8990	0.8983	0.8979	0.8976	0.8974
	4	CW	0.9013	0.9007	0.9003	0.9001	0.9000
		LS	0.9002	0.8996	0.8992	0.8990	0.8989
	5	CW	0.9011	0.9005	0.9002	0.9001	0.9000
		LS	0.9011	0.9005	0.9002	0.9001	0.9000
	6	CW	0.9008	0.9004	0.9001	0.9000	0.9000
		LS	0.9017	0.9012	0.9010	0.9009	0.9008
	7	CW	0.9007	0.9003	0.9001	0.9000	0.9000
		LS	0.9021	0.9017	0.9015	0.9014	0.9014
0.95	3	CW	0.9504	0.9502	0.9501	0.9500	0.9500
		LS	0.9497	0.9495	0.9494	0.9494	0.9494
	4	CW	0.9497	0.9495	0.9494	0.9494	0.9494
		LS	0.9500	0.9498	0.9498	0.9497	0.9497
	5	CW	0.9502	0.9501	0.9500	0.9500	0.9500
		LS	0.9502	0.9501	0.9500	0.9500	0.9500
	6	CW	0.9502	0.9501	0.9500	0.9500	0.9500
		LS	0.9504	0.9503	0.9502	0.9502	0.9502
	7	CW	0.9502	0.9501	0.9500	0.9500	0.9500
		LS	0.9505	0.9504	0.9504	0.9504	0.9504

Finally, note that, when tied observations are present, the midranks are suggested to replace the associated ranks in (2.1)-(2.3). In fact, this implies that we replace $P(X_i < X_{i+1})$ by $P(X_i < X_{i+1}) + 0.5P(X_i = X_{i+1})$ for the relative effect between the ith and the $(i+1)$th groups.

3. Monte Carlo Study

A simulation study is conducted for studying the coverage probability of the proposed simultaneous lower confidence bounds, denoted by CW, for a variety of distributions. We also investigate the one suggested in Lee and Spurrier [9], denoted by LS, for comparing adjacent location parameters on holding its confidence level when the underlying distributions are different not only in location parameters, but also in scale parameters.

Two different types of distributions are under study: (1) Normal: $X_i \sim N(\theta_i, \sigma_i)$, $i = 1, ..., k$, and (2) Exponential: $X_i \sim E(\theta_i, \sigma_i)$, $i = 1, ..., k$. We consider $k = 4$ groups each with a common sample size of $n_1 = \cdots = n_4 = n = 10, 20$ or 30. The IMSL (International Mathematical and Statistical Library) random number generators DRNNOR and DRNEXP are used to generate normal and exponential variates, respectively. For each set of distributions, the p_i's and the differences of the successive population medians are computed. Under the nominal level $1-\alpha = 0.95$, the CW bounds are constructed for the p_i's and the LS bounds are established for the adjacent-median differences. We perform the simulation with 2000 replications for each set of distributions and samples. The proportion of the confidence bounds which cover the true probabilities or parameters is then used to estimate the coverage probability of the confidence bounds. Finally, the results are reported in Tables 2 and 3. Note that the standard error of the estimate is about 0.005 $(= \sqrt{(0.95)(0.05)/2000})$.

Table 2. Simulated coverages for simultaneous lower confidence bounds under nominal level 0.95 and $n_1 = \ldots = n_4 = n$ when the underlying distributions are normal

$(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$		(2, 2, 2, 2)		(5, 2, 2, 2)		(10, 5, 2, 2)	
(p_1, p_2, p_3)	Method	$n = 10$	20	10	20	10	20
(0.5, 0.5, 0.5)	CW	0.951	0.949	0.952	0.948	0.948	0.945
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.5, 0.5, 0.6)	CW	0.952	0.952	0.953	0.951	0.949	0.948
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.5, 0.7, 0.5)	CW	0.945	0.947	0.946	0.946	0942	0.947
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.7, 0.5, 0.5)	CW	0.949	0.951	0.946	0.947	0.947	0.947
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.5, 0.6, 0.7)	CW	0.951	0.951	0.952	0.951	0.950	0.951
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.6, 0.5, 0.6)	CW	0.953	0.951	0.956	0.951	0.951	0.950
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.7, 0.7, 0.5)	CW	0.943	0.949	0.939 –	0.945	0.941	0.949
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.8, 0.6, 0.5)	CW	0.951	0.951	0.935 –	0.946	0.943	0.945
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.7, 0.6, 0.6)	CW	0.951	0.954	0.948	0.950	0.951	0.953
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –
(0.6, 0.7, 0.8)	CW	0.945	0.945	0.947	0.945	0.942	0.947
	LS	0.951	0.947	0.942	0.941	0.940	0.937 –

±: At least two standard deviations above (below) $1 - \alpha = 0.95$.

Table 3. Simulated coverages for simultaneous lower confidence bounds under nominal level 0.95 and $n_1 = \ldots = n_4 = n$ when the underlying distributions are normal

$(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$		(2, 2, 2, 2)		(5, 2, 2, 2)		(10, 5, 2, 2)	
(p_1, p_2, p_3)	method	$n = 10$	20	10	20	10	20
(0.5, 0.5, 0.5)	CW	0.955	0.955	0.956	0.953	0.953	0.953
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.5, 0.5, 0.6)	CW	0.955	0.956	0.956	0.955	0.954	0.955
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.5, 0.7, 0.5)	CW	0.951	0.957	0.952	0.955	0.951	0.953
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.7, 0.5, 0.5)	CW	0.950	0.958	0.944	0.958	0.943	0.957
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.5, 0.6, 0.7)	CW	0.943	0.959	0.944	0.958	0.946	0.958
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.6, 0.5, 0.6)	CW	0.954	0.957	0.956	0.958	0.951	0.959
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.7, 0.7, 0.5)	CW	0.946	0.959	0.940	0.960	0.941	0.957
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.8, 0.6, 0.5)	CW	0.931−	0.956	0.930−	0.956	0.933−	0.957
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.7, 0.6, 0.6)	CW	0.948	0.961+	0.942	0.961+	0.946	0.960
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−
(0.6, 0.7, 0.8)	CW	0.929−	0.959	0.930−	0.959	0.927−	0.959
	LS	0.957	0.951	0.940	0.922−	0.924−	0.903−

±: At least two standard deviations above (below) $1 - \alpha = 0.95$.

The coverage probabilities presented in Tables 2 and 3 indicate that the CW procedure generally holds its nominal confidence level except for the cases with small sample sizes as 10 and some of the p_i's as high as 0.7 or 0.8. The LS procedure also reasonably maintains its nominal

confidence level when the underlying distributions differ only in location parameters. However, when the scale parameters of the underlying distributions are widely different, the coverage probability of the LS procedure does not reach the nominal level 0.95.

4. An Example

To investigate the effect of a chemical, called *vinylidene fluoride*, on liver damage, the National Toxicology Program conducted a subchronic animal study in which groups of ten male Fischer 344 rats were assigned to receive by inhalation exposure one of six dosages of vinylidene fluoride and three serum enzymes were measured for each rat. In this section, we consider the serum enzyme, namely, SDH, under the four dosages (parts/million) of vinylidene fluoride: 0, 1500, 5000 and 15000 (Dietz [3]).

We explore the assumptions of normal distribution and equal variation for the data. Note that the box plot, displayed in Figure 1, reveals that the median SDH is non-decreasing with the dosage of vinylidene fluoride, while the variation of SDH in the zero-dosage group seems to be larger than the ones in the other groups. The Lilliefors normal test further gives the p-value of 0.038, 0.50, 0.014 and 0.50 for the four groups under study. Hence, we learned that the normal distribution assumption is not feasible for the first and the third groups. Moreover, the underlying distributions may not be different only in location parameters. Therefore, we employ the CW procedure to analyze the SDH data.

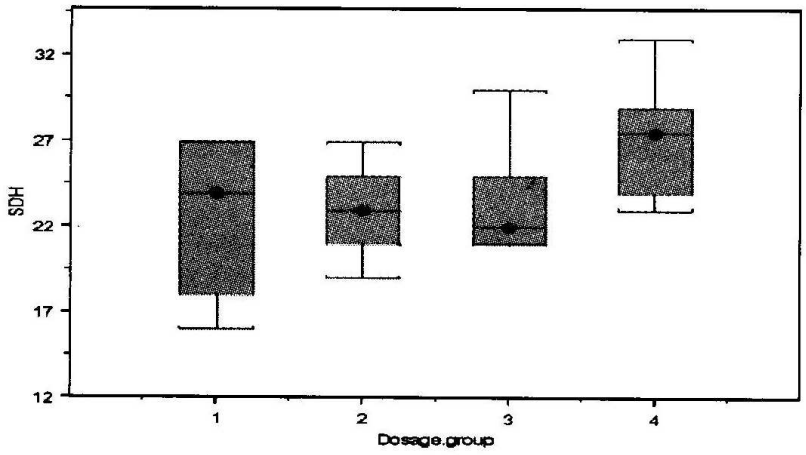

Figure 1. Box plot for seurm enzyme SDH.

Note, as indicated in Dietz [3], that the SDH is, of statistical significance, increasing with the dosage of vinylidene fluoride. The 95% lower confidence bounds for (p_1, p_2, p_3) proposed herein are further obtained as (0.21, 0.30, 0.51). Therefore, at 95% confidence level, we claim that the first dosage of vinylidene fluoride may be regarded as unharmful as the zero-dose control to the liver damage. The dosages of 1500 and 5000 are equivalent in terms of the SDH value. However, the dosage 15000 preserves a significant effect on causing larger SDH value than the dosage 5000. Finally, the data analysis suggests only the highest dosage of vinylidene fluoride under study is significantly bringing damage to the liver of rat than the other lower dosages of vinylidene fluoride.

References

[1] P. Billingsley, Probability and Measure, 2nd ed., Wiley, New York, 1991.

[2] Y. I. Chen and D. A. Wolfe, Modifications of the Mack-Wolfe umbrella tests for a generalized Behrens-Fisher problem, Canad. J. Statist. 18 (1990), 245-253.

[3] E. J. Dietz, Multivariate generalizations of Jonckheere's test for ordered alternatives, Comm. Statist. Theory Methods 18 (1989), 3763-3783.

[4] Z. Drezner, Computation of the multivariate normal integral, ACM Trans. Math. Software 18 (1992), 470-480.

[5] B. K. Ghosh, A comparison of some approximate confidence intervals for the binomial parameter, J. Amer. Statist. Assoc. 74 (1979), 894-900.

[6] M. Halperin et al., Distribution-free confidence intervals for $\Pr(X_1 < X_2)$, Biometrics 43 (1987), 71-80.

[7] D. Hunter, An upper bound for the probability of a union, J. Appl. Probab. 13 (1976), 597-603.

[8] R. E. Lee and J. D. Spurrier, Successive comparisons between ordered treatments, J. Statist. Plann. Inference 43 (1995a), 323-330.

[9] R. E. Lee and J. D. Spurrier, Distribution-free multiple comparisons between successive treatments, Nonparametric Statist. 5 (1995b), 261-273.

[10] H. W. Lilliefors, On the Komogorov-Smirnov test for normality with mean and variance unknown, J. Amer. Statist. Assoc. 62 (1967), 399-402.

[11] W. Liu et al., Simultaneous confidence interval estimation for successive comparisons of ordered treatment effects, J. Statist. Plann. Inference 88 (2000), 75-86.

[12] H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Statist. 18 (1947), 50-60.

[13] R. W. Mee, Confidence intervals for probabilities and tolerance regions based on a generalization of the Mann-Whitney statistic, J. Amer. Statist. Assoc. 85 (1990), 793-800.

[14] P. K. Sen, A note on asymptotically distribution-free confidence bounds for $\Pr(X < Y)$ based on two independent samples, Sankhyã, Ser. A 29 (1967), 95-102.

[15] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), 80-83.

[16] K. J. Worsley, An improved Bonferroni inequality and applications, Biometrika 69 (1982), 297-302.

■