Finding the Maximum Safe Dose Level for Heteroscedastic Data

Ajit C. Tamhane
Departments of IE/MS and of Statistics
2145 Sheridan Road
Northwestern University
Evanston, IL 60208

Brent R. Logan
Division of Biostatistics
Medical College of Wisconsin
Watertown Plank Rd.
Milwaukee, WI 53226

May 13, 2003

Abstract

In this paper we extend to the heteroscedastic setting the multiple stepwise test procedures proposed in Tamhane, Dunnett, Green and Wetherington (2001) for finding the maximum safe dose (MAXSD). Toxicological data are often heteroscedastic, and therefore the extensions given herein should be highly useful in practice. Simulations are performed to study the type I familywise error rate (FWE) and power properties of the procedures. A real data example is given to illustrate the procedures.

Keywords and Phrases: Dose response; Multiple comparisons; Toxicology; Unequal variances

1. Introduction

In Tamhane et al. (2001) (abbreviated as TDGW hereafter) we developed three stepwise multiple testing procedures (labeled SD1PC, SD2PC and SD1HC) for finding the maximum safe dose (MAXSD) level of a compound using toxicological data. These procedures are applicable to homoscedastic data. However, toxicological data are often heteroscedastic, i.e., the data from different doses have different variances. In this paper we extend the procedures in TDGW to the heteroscedastic setting.

The outline of the paper is as follows. Section 2 gives the problem formulation and defines the notation. Section 3 states the heteroscedastic extensions of the SD1PC, SD2PC and SD1HC procedures. Section 4 describes the simulation results on the type I error rates and powers of SD1PC, SD2PC and SD1HC. A real data example is given in Section 5. Finally, conclusions are summarized in Section 6.

2. Preliminaries and Notation

Let $0, 1, \ldots, k$ denote increasing dose levels with 0 being the zero dose level (control). Assume that the measurements y_{ij} $(j = 1, 2, \ldots, n_i)$ on dose i are independent and normally distributed with mean μ_i and variance σ_i^2 (denoted by $y_{ij} \sim N(\mu_i, \sigma_i^2)$). Let $\overline{y}_i$ denote the sample means and s_i^2 denote the sample variances based on $n_i - 1$ d.f., $i = 0, 1, \ldots, k$. A small μ_i (e.g., a lower yield of a crop contaminated by a herbicide) implies a more toxic response. The case where a larger μ_i represents a more toxic response can be handled analogously.

We regard as *unsafe* a decrease in mean yield below a specified percentage (e.g., 10%) of the mean yield μ_0 at the zero dose level. In general, we specify $\lambda < 1$ (e.g., $\lambda = 0.90$ for a 10% decrease in the mean yield compared to μ_0). All doses with $\mu_i \leq \lambda\mu_0$ are regarded as unsafe and those with $\mu_i > \lambda\mu_0$ are regarded as safe. The maximum safe dose (MAXSD) for specified λ is defined as

$$\text{MAXSD} = \max\{i : \mu_j > \lambda\mu_0 \ \ \forall j \leq i\}. \tag{2.1}$$

Define the minimum unsafe dose (MINUD) by MINUD = MAXSD + 1. If a larger μ_i represents a more toxic response then $\lambda > 1$. In that case, doses with $\mu_i \geq \lambda\mu_0$ are regarded

as unsafe and those with $\mu_i < \lambda\mu_0$ are regarded as safe. The MAXSD is then defined as $\text{MAXSD} = \max\{i : \mu_j < \lambda\mu_0 \ \ \forall j \leq i\}$.

We want to guarantee that the probability that *any* unsafe dose is declared safe is no more than a specified constant α. If $\widehat{\text{MAXSD}}$ denotes the estimated MAXSD then this requirement translates to

$$P\{\widehat{\text{MAXSD}} > \text{MAXSD}\} \leq \alpha. \tag{2.2}$$

Now consider the family of hypothesis testing problems:

$$H_{0i} : \mu_i \leq \lambda\mu_0 \text{ vs. } H_{1i} : \mu_i > \lambda\mu_0 \ \ (1 \leq i \leq k). \tag{2.3}$$

Here H_{0i} states that the i-th dose is unsafe and H_{1i} states that the i-th dose is safe. After testing the hypotheses we set $\widehat{\text{MAXSD}} = \max\{i : H_{0i} \text{ is rejected}\}$. The error probability requirement (2.2) is satisfied if we control the type I FWE at level α for the family of hypotheses (2.3):

$$\text{FWE} = P\{\text{Any true } H_{0i} \text{ is rejected}\} \leq \alpha. \tag{2.4}$$

It was shown in Section 8 of TDGW that this requirement is satisfied if the null hypotheses in (2.3) are tested in a step-down (SD) manner beginning with H_{01}; if it is rejected then test H_{02} and so on, each at level α, which is what SD2PC does. On the other hand, SD1PC and SD1HC test $\bigcap_{j=i}^{k} H_{0j} \subseteq H_{0i}$, in a step-down manner each at level α. Under the assumption of monotonicity: $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_k$, we have $H_{0i} = \bigcap_{j=i}^{k} H_{0j}$, and therefore SD1PC and SD1HC procedures also control (2.2). However, if the means are not monotone then only SD2PC controls this requirement as shown by Bauer (1997).

3. Test Procedures

SD1PC and SD2PC are based on pairwise contrasts and are described in Section 3.1. SD1HC is based on Helmert contrasts and is described in Section 3.2.

3.1 SD1PC and SD2PC

Define the t-statistics for pairwise contrasts as

$$t_i = \frac{\overline{y}_i - \lambda\overline{y}_0}{\sqrt{s_i^2/n_i + \lambda^2 s_0^2/n_0}} \ \ (1 \leq i \leq k). \tag{3.1}$$

The SD1PC procedure tests H_{0i} using the statistic

$$t_{\max,i} = \max_{i \le j \le k} t_j, \tag{3.2}$$

while SD2PC uses ordinary t-tests based on the t_i. Each procedure tests H_{0i} if H_{0j} for $j = 1, \ldots, i-1$ have been rejected and rejects H_{0i} if the observed test statistic exceeds the upper α critical point of its null distribution. Equivalently, these tests can be applied in terms of the p-values. If p_i denotes the marginal p-value associated with t_i then SD2PC rejects H_{0i} if $p_i < \alpha$. To apply SD1PC one needs to calculate the adjusted p-value, denoted by $\tilde{p}_i$, which is the probability that under H_{0i}, the statistic $t_{\max,i}$ is at least as large as its observed value. To calculate $\tilde{p}_i$, we need to know the joint distribution of t_j $(i \le j \le k)$.

The exact joint distribution of the t_i are intractable and depends on the unknown variances σ_i^2. Therefore we must use approximations to the critical points of these distributions. The marginal distribution of each t_i can be approximated by Student's t-distribution with d.f. obtained using the Welch- Satterthwaite approximation:

$$\nu_i = \frac{(\lambda^2 s_0^2/n_0 + s_i^2/n_i)^2}{\lambda^4 s_0^4/n_0^2(n_0-1) + s_i^4/n_i^2(n_i-1)}. \tag{3.3}$$

For SD1PC we need the upper α critical point of $t_{\max,i}$, which involves the joint distribution of $t_i, t_{i+1}, \ldots, t_k$. This distribution can be approximated using the $(k-i+1)$-variate t-distribution for $i = 1, 2, \ldots, k$. The correlation matrix of this distribution can be estimated for $i \le j < j' \le k$ using

$$\rho_{jj'} = \frac{\lambda^2 s_0^2/n_0}{\sqrt{(\lambda^2 s_0^2/n_0 + s_j^2/n_j)(\lambda^2 s_0^2/n_0 + s_{j'}^2/n_{j'})}}. \tag{3.4}$$

These correlations have the product form $\rho_{jj'} = \gamma_j \gamma_{j'}$ where

$$\gamma_j = \sqrt{\frac{\lambda^2 s_0^2/n_0}{\lambda^2 s_0^2/n_0 + s_j^2/n_j}}. \tag{3.5}$$

So the computation of the desired critical point can be performed efficiently using the iterated integral representation given in equation (1.1a) in Appendix 3 of Hochberg and Tamhane (1987). A Fortran program for this purpose based on Dunnett (1989) is available from `http://lib.stat.cmu.edu/general`. The average of the $\rho_{jj'}$ can be used to yield a good

approximation; see Hochberg and Tamhane (1987, p. 146). For the d.f. of the multivariate t-distribution, the average of the ν_j $(i \le j \le k)$ can be used.

3.2 SD1HC Procedure

A Helmert contrast is defined as

$$C_{ij} = (\overline{y}_i + \cdots + \overline{y}_j) - (j - i + 1)\lambda\overline{y}_0 \quad (1 \le j \le i). \tag{3.6}$$

The standard error of C_{ij} equals

$$s.e.(C_{ij}) = \sqrt{(j - i + 1)^2\lambda^2 s_0^2/n_0 + \sum_{h=i}^{j}(s_h^2/n_h)}. \tag{3.7}$$

Define the statistic

$$t_{ij} = \frac{C_{ij}}{s.e.(C_{ij})}. \tag{3.8}$$

Then in analogy with (3.2), the test statistic for testing H_{0i} is

$$t_{\max,i} = \max_{i \le j \le k} t_{ij}.$$

As in the case of SD1PC, H_{0i} is tested if H_{0j}, $j = 1, \ldots, i-1$ are rejected and if the adjusted p-value associated with $t_{\max,i}$, denoted by $\widetilde{p}_i$, is less than α. To calculate $\widetilde{p}_i$, we need to know the joint distribution of t_{ij} $(i \le j \le k)$.

The marginal distribution of t_{ij} can be approximated by Student's t-distribution with the Welch-Satterthwaite d.f.:

$$\nu_{ij} = \frac{\{(j - i + 1)^2\lambda^2 s_0^2/n_0 + \sum_{h=i}^{j}(s_h^2/n_h)\}^2}{\{(j - i + 1)^4\lambda^4 s_0^4/n_0^2(n_0 - 1) + \sum_{h=i}^{j}[s_h^4/n_h^2(n_h - 1)]\}}. \tag{3.9}$$

The correlation matrix of t_{ij} for $i \le j < j' \le k$ can be estimated using

$$\rho_{jj'}^{(i)} = \frac{\sum_{h=i}^{j}(s_h^2/n_h) + (j - i + 1)(j' - i + 1)\lambda^2 s_0^2/n_0}{\sqrt{[\sum_{h=i}^{j}(s_h^2/n_h) + (j - i + 1)^2\lambda^2 s_0^2/n_0][\sum_{h=i}^{j'}(s_h^2/n_h) + (j' - i + 1)^2\lambda^2 s_0^2/n_0]}}. \tag{3.10}$$

These correlations do not have the product structure. The SAS-IML program based on Genz and Bretz (1999) with arbitrary correlation matrices available from `www.bioinf.uni-hannover.de` can be used in this case. The degrees of freedom for the multivariate t can be approximated by the average of the ν_{ij} $(i \le j \le k)$.

4. Simulation Results

We performed Monte Carlo simulations to study the FWEs and powers of SD1PC, SD2PC and SD1HC. We studied two types of configurations: (i) constant coefficient of variation (CV) in which the σ_i are proportional to the μ_i, and (ii) constant variance among the positive doses (which is different from the zero dose variance). The μ_i configurations are the same as those in Table 5 of TDGW except that when studying the FWE, we set μ_{MINUD} equal to $\lambda\mu_0$ since it is the least favorable configuration at which the max FWE is attained. The results for the constant CV case are given in Table 3 (FWE) and Table 4 (Power). The results for the constant variance among the positive doses case are given in Table 5 (FWE) and Table 6 (Power). All estimates are based on 10,000 simulation runs.

It is seen from Tables 3 and 5 that the approximate SD1PC, SD2PC and SD1HC procedures control the FWE at or below $\alpha = 0.05$ for all the configurations studied. The FWE exceeds the nominal $\alpha = 0.05$ at the 5% significance level if the estimated FWE exceeds

$$0.05 + 1.96\sqrt{\frac{(0.05)(0.95)}{10,000}} = 0.054.$$

There are only a few cases where the estimated FWE is > 0.054, and the exceedence is not practically significant in these cases. In many linear configurations the FWEs of SD1PC and SD2PC are quite small and zero in several cases. This is because those configurations are not least favorable in terms of maximizing the FWE. In conclusion, the approximations are quite accurate, but a bit conservative.

We now turn to Tables 4 and 6 for the powers of SD1PC, SD2PC and SD1HC. The highest of the three powers is marked with an asterisk in each case. We see that SD1PC always has less power than either SD2PC or SD1HC (except in two cases, where its power is either equal to or slightly higher than that of its closest competitor); thus SD1PC is not recommended in any situation. Neither SD2PC nor SD1HC dominates the other procedure for step configurations, but for linear configurations SD2PC is uniformly superior. These results are consonant with those in TDGW.

5. Example

A 90-day routine rat study was conducted to evaluate the toxicity of a crop protection

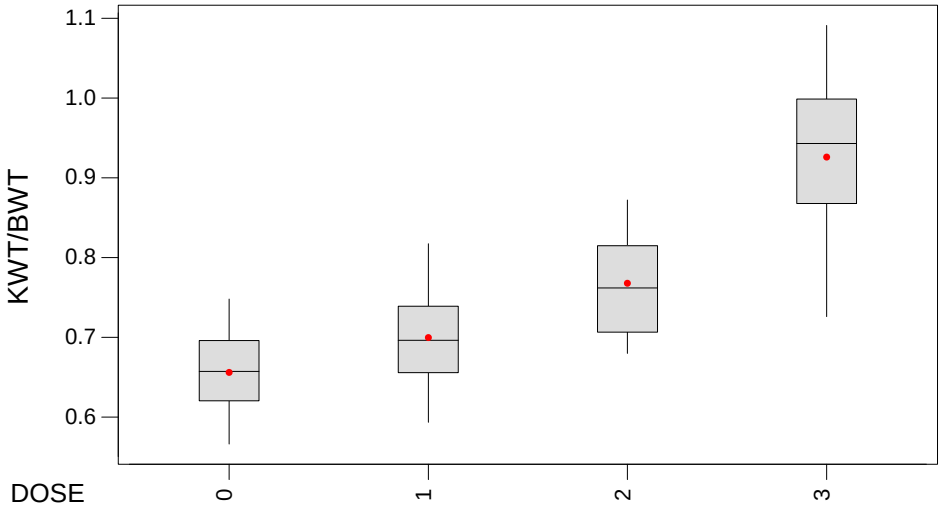

Figure 1: Box Plot of Kidney Data

compound. Test substance was added directly to the rodent diet and was thoroughly mixed to ensure homogeneous distribution. Three doses of the compound were compared with a zero dose control. The sample sizes in the four groups were $n_0 = 18, n_1 = 20, n_2 = 19$ and $n_3 = 18$. The variable of interest was the kidney weight to the body weight ratio. A large value of this ratio is regarded as unsafe with a threshold of a 15% average increase over its value for the zero dose. Thus we specify $\lambda = 1.15$.

The data are given in Table 1. The box plot of the data is shown in Figure 1. The dose means seem to increase quadratically. The mean differences are highly significant ($F = 52.82$ with $p = 0.000$). The within group variances also appear to be different across the dose groups (Bartlett's $\chi^2 = 10.446$ with $p = 0.015$, Levene's $F = 3.115$ with $p = 0.030$). However, the normality of the data does not seem to be in serious doubt as can be seen from the normal plot of the residuals in Figure 2. There appear to be a few outliers in this plot, but they are not identified as such by outlier tests; as a result, they are not deleted. Thus the assumptions necessary to apply the proposed methods seem to be satisfied.

The inequalities in the hypotheses (2.3) are reversed as follows:

$$H_{0i} : \mu_i \geq \lambda\mu_0 \text{ vs. } H_{1i} : \mu_i < \lambda\mu_0, \quad i = 1, 2, \ldots, k.$$

Table 1: (Kidney Wt./Body Wt.)$\times 10^3$

Dose			
0	1	2	3
6.593	7.062	7.001	9.569
7.480	7.347	8.706	9.362
6.930	7.733	7.257	1.091
5.662	7.396	7.743	0.996
6.789	8.173	7.026	0.950
7.268	6.938	8.561	0.991
6.647	6.988	7.674	0.854
6.443	6.621	7.450	1.040
6.713	7.508	8.188	1.042
6.057	6.657	8.150	1.006
6.253	7.787	7.619	0.967
7.045	6.537	8.722	0.819
6.552	7.369	7.387	0.899
5.668	6.623	6.798	0.735
6.354	6.456	7.617	0.726
6.511	6.507	8.071	0.902
7.111	6.154	7.020	0.885
6.015	5.934	7.821	0.872
	6.909	7.063	
	7.252		
n_i	18	20	19	18
Mean	6.5606	6.9975	7.6778	9.2606
SD	0.5094	0.5755	0.5949	1.0052

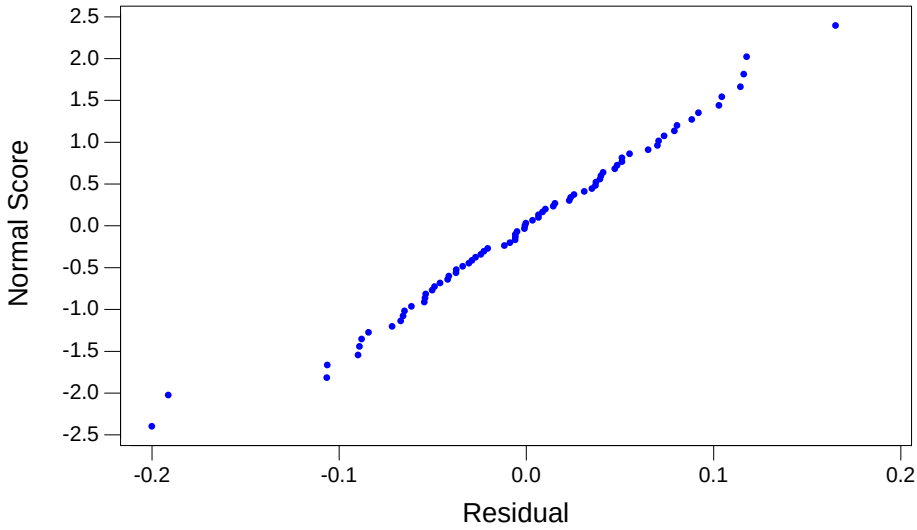

Figure 2: Normal Plot of the Residuals for Kidney Data

Also, the sign of the test statistic (3.1) must be reversed. In other words, the pairwise contrast t-statistics are defined as

$$t_i = \frac{\lambda \overline{y}_0 - \overline{y}_i}{\sqrt{\lambda^2 s_0^2/n_0 + s_i^2/n_i}} \quad (1 \leq i \leq k).$$

Similarly, the signs on the Helmert contrasts defined in (3.6) must also be reversed. After making these minor modifications we are now ready to apply the procedures.

SD1PC and SD2PC Procedures: Table 2 lists the t-statistics and their d.f. computed using (3.1) and (3.3), respectively, and their marginal p-values. We first apply SD2PC. Since $p_1 = 0.003 < \alpha = 0.05$, we reject H_{01} and conclude that dose 1 is safe. But $p_2 = 0.751 > \alpha = 0.05$, so H_{02} can't be rejected and testing stops with $\widehat{\text{MAXSD}} = 1$.

To apply SD1PC, we need to calculate $\tilde{p}_1$ corresponding to $\max_{1 \leq i \leq 3} t_i = t_1 = 2.899$ for which we need the correlation matrix of t_1, t_2, t_3. This matrix is as follows (calculated using (3.4)):

$$\begin{bmatrix} 1 & 0.552 & 0.537 \\ & 1 & 0.520 \\ & & 1 \end{bmatrix}.$$

The average d.f. equals 32.58, which is rounded down to 32. Using these values we get

Table 2: t-Statistics for Pairwise Contrasts, Their Degrees of Freedom and p-Values

	Dose		
i	1	2	3
t_i	2.899	-0.686	-6.257
ν_i	35.44	34.94	27.35
p_i	0.003	0.751	1.000

$\widetilde{p}_1 = 0.009 < \alpha = 0.05$, so we reject H_{01} and conclude that dose 1 is safe. There is no need to calculate $\widetilde{p}_2$ since it will be greater than $p_2 = 0.751$ and hence greater than $\alpha = 0.05$, so testing stops with $\widehat{\text{MAXSD}} = 1$.

SD1HC Procedure: We begin by testing H_{01}. The three Helmert contrast t-statistics with their d.f. are (computed using (3.8) and (3.9)):

$$t_{11} = 2.899, t_{12} = 1.240, t_{13} = -2.539; \nu_{11} = 35.44, \nu_{12} = 33.05, \nu_{13} = 35.43.$$

The average d.f. equals 34.64, which is rounded down to 34. The correlation matrix (computed using (3.10)) is

$$\begin{bmatrix} 1 & 0.868 & 0.762 \\ & 1 & 0.874 \\ & & 1 \end{bmatrix}.$$

Notice that these correlations are higher than those for pairwise contrasts. As a result, the adjusted p-value, $\widetilde{p}_1$, corresponding to $\max_{1 \le j \le 3} t_{1j} = t_{11} = 2.899$ is smaller, namely, $\widetilde{p}_1 = 0.007 < \alpha = 0.05$. Thus H_{01} is rejected. The t-statistics and their d.f. for testing H_{02} are

$$t_{22} = -0.686, t_{23} = -4.758; \nu_{22} = 34.94, \nu_{23} = 41.72.$$

Clearly, t_{22} is nonsignificant, so testing stops with $\widehat{\text{MAXSD}} = 1$.

It is worth noting that if we use $\lambda = 1.10$ then $p_1 = 0.121$ for SD2PC, $\widetilde{p}_1 = 0.254$ for SD1PC and $\widetilde{p}_1 = 0.196$ for SD1HC. Thus none of the procedures is able to find dose 1 as MAXSD.

6. Concluding Remarks

We have given heteroscedastic extensions of SD1PC, SD2PC and SD1HC for identifying MAXSD in dose response studies for safety testing. All three procedures control the type I error rates. However, in terms of power, SD1PC is dominated by either SD2PC or SD1HC as in the homoscedastic case considered in TDGW. So either SD2PC or SD1HC is preferred, the former in the linear dose response case and the latter in the step dose response case. SD2PC has the advantage of being simpler and more generally applicable, but its power can be quite low for the step dose response case if MAXSD is a high dose.

Acknowledgments: This research was partially supported by an educational grant from DuPont. We thank Dr. John Green of DuPont Safety, Health and Environmental Excellence Center, Haskell Laboratory for Toxicology and Industrial Medicine for providing the data for the problem and helpful discussions.

References

1. Dunnett, C. W. (1989), "Multivariate normal probability integrals with product correlation structure," Algorithm AS251, *Applied Statistics*, **38**, 564–579.

2. Genz, A. and Bretz, F. (1999), "Numerical computation of multivariate t-probabilities with application to power calculation of multiple contrasts," *Journal of Statistical Computation and Simulation*, **63**, 361–378.

3. Hochberg, Y. and Tamhane, A. C. (1987), *Multiple Comparison Procedures*, New York: Wiley.

4. Marcus, R., Peritz, E. and Gabriel, K. R. (1976), "On closed testing procedures with special reference to ordered analysis of variance," *Biometrika*, **63**, 655–660.

5. Tamhane, A. C., Hochberg, Y. and Dunnett, C. W. (1996), "Multiple test procedures for dose finding," *Biometrics*, **52**, 21–37.

6. Tamhane, A. C., Dunnett, C. W., Green, J. W. and Wetherington, J. F. (2001), "Multiple test procedures for identifying a safe dose," *Journal of the American Statistical Association*, (2001), **96**, 835–843.

7. Westfall, P. and Young, S. S. (1993), *Resampling-Based Multiple Testing*, New York, John Wiley.

Table 3: Simulated Type I Error Rates of SD1PC, SD2PC and SD1HC Under Constant CV When $\lambda = 0.85$, $\mu_0 = 20$, $\mu_{\text{MAXSD}} = 18$, $\sigma_0 = 1$, $n_0 = 16$, $n_1 = \cdots = n_5 = 8$ and $\alpha = 0.05$.

Configuration	MAXSD	$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$	SD1PC	SD2PC	SD1HC
Step	1	(18,17,17,17,17)	0.053	0.045	0.054
	2	(18,18,17,17,17)	0.054	0.041	0.051
	2	(20,18,17,17,17)	0.048	0.048	0.053
	3	(18,18,18,17,17)	0.046	0.039	0.050
	3	(20,18,18,17,17)	0.046	0.042	0.053
	3	(20,20,18,17,17)	0.047	0.046	0.053
	4	(18,18,18,18,17)	0.044	0.036	0.048
	4	(20,18,18,18,17)	0.047	0.041	0.051
	4	(20,20,18,18,17)	0.047	0.045	0.050
	4	(20,20,20,18,17)	0.044	0.044	0.048
Linear	1	(18,14,10,6,2)	0.000	0.000	0.000
	1	(18,17,15,13,11)	0.014	0.044	0.028
	2	(19,18,17,15,13)	0.018	0.047	0.031
	2	(19,18,14,10,6)	0.000	0.000	0.000
	3	(19,19,18,17,15)	0.027	0.044	0.038
	3	(19,19,18,14,10)	0.000	0.000	0.000
	4	(19,19,19,18,17)	0.045	0.044	0.046
	4	(19,19,19,18,14)	0.000	0.000	0.000

Table 4: Simulated Powers of SD1PC, SD2PC and SD1HC for Identifying MAXSD Under Constant CV When $\lambda = 0.85$, $\mu_0 = 20$, $\mu_{\text{MAXSD}} = 18$, $\sigma_0 = 1$, $n_0 = 16$, $n_1 = \cdots = n_5 = 8$ and $\alpha = 0.05$.

Configuration	MAXSD	$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$	SD1PC	SD2PC	SD1HC
Step	1	(18,17,17,17,17)	0.519	0.753*	0.653
	2	(18,18,17,17,17)	0.530	0.615	0.657*
	2	(20,18,17,17,17)	0.544	0.745*	0.673
	3	(18,18,18,17,17)	0.563	0.516	0.661*
	3	(20,18,18,17,17)	0.571	0.615	0.672*
	3	(20,20,18,17,17)	0.593	0.750*	0.686
	4	(18,18,18,18,17)	0.602	0.436	0.681*
	4	(20,18,18,18,17)	0.616	0.509	0.686*
	4	(20,20,18,18,17)	0.625	0.610	0.692*
	4	(20,20,20,18,17)	0.649	0.744*	0.709
	5	(18,18,18,18,18)	0.727	0.417	0.764*
	5	(20,18,18,18,18)	0.740	0.490	0.766*
	5	(20,20,18,18,18)	0.741	0.561	0.765*
	5	(20,20,20,18,18)	0.757	0.657	0.780*
	5	(20,20,20,20,18)	0.801*	0.800	0.797
Linear	1	(18,14,10,6,2)	0.593	0.799*	0.717
	1	(18,17,15,13,11)	0.555	0.748*	0.683
	2	(19,18,17,15,13)	0.575	0.738*	0.692
	2	(19,18,14,10,6)	0.615	0.793*	0.726
	3	(19,19,18,17,15)	0.619	0.749*	0.702
	3	(19,19,18,14,10)	0.651	0.798*	0.736
	4	(19,19,19,18,17)	0.649	0.749*	0.705
	4	(19,19,19,18,14)	0.694	0.791*	0.755

Table 5: Simulated Type I Error Rates of SD1PC, SD2PC and SD1HC When $\lambda = 0.85$, $\mu_0 = 20$, $\mu_{\text{MAXSD}} = 18$, $\sigma_0 = 0.5$, $\sigma_1 = \cdots = \sigma_5 = 2$, $n_0 = 16$, $n_1 = \cdots = n_5 = 8$ and $\alpha = 0.05$.

Configuration	MAXSD	$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$	SD1PC	SD2PC	SD1HC
Step	1	(18,17,17,17,17)	0.050	0.043	0.054
	2	(18,18,17,17,17)	0.051	0.036	0.053
	2	(20,18,17,17,17)	0.054	0.041	0.050
	3	(18,18,18,17,17)	0.048	0.033	0.048
	3	(20,18,18,17,17)	0.050	0.036	0.056
	3	(20,20,18,17,17)	0.046	0.044	0.055
	4	(18,18,18,18,17)	0.039	0.025	0.053
	4	(20,18,18,18,17)	0.041	0.030	0.054
	4	(20,20,18,18,17)	0.042	0.033	0.047
	4	(20,20,20,18,17)	0.049	0.045	0.046
Linear	1	(18,14,10,6,2)	0.000	0.000	0.000
	1	(18,17,15,13,11)	0.014	0.040	0.030
	2	(19,18,17,15,13)	0.019	0.047	0.030
	2	(19,18,14,10,6)	0.000	0.000	0.000
	3	(19,19,18,17,15)	0.028	0.045	0.039
	3	(19,19,18,14,10)	0.000	0.000	0.000
	4	(19,19,19,18,17)	0.046	0.044	0.049
	4	(19,19,19,18,14)	0.000	0.000	0.000

Table 6: Simulated Powers of SD1PC, SD2PC and SD1HC for Identifying MAXSD When $\lambda = 0.85$, $\mu_0 = 20$, $\mu_{\text{MAXSD}} = 18$,$\sigma_0 = 0.5$, $\sigma_1 = \cdots = \sigma_5 = 2$, $n_0 = 16$, $n_1 = \cdots = n_5 = 8$ and $\alpha = 0.05$.

Configuration	MAXSD	$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$	SD1PC	SD2PC	SD1HC
Step	1	(18,17,17,17,17)	0.486	0.756*	0.682
	2	(18,18,17,17,17)	0.497	0.602	0.682*
	2	(20,18,17,17,17)	0.518	0.757*	0.688
	3	(18,18,18,17,17)	0.539	0.482	0.693*
	3	(20,18,18,17,17)	0.552	0.594	0.691*
	3	(20,20,18,17,17)	0.579	0.750*	0.692
	4	(18,18,18,18,17)	0.593	0.395	0.701*
	4	(20,18,18,18,17)	0.609	0.494	0.717*
	4	(20,20,18,18,17)	0.613	0.598	0.711*
	4	(20,20,20,18,17)	0.642	0.746*	0.726
	5	(18,18,18,18,18)	0.725	0.345	0.799*
	5	(20,18,18,18,18)	0.733	0.421	0.785*
	5	(20,20,18,18,18)	0.749	0.514	0.783*
	5	(20,20,20,18,18)	0.765	0.647	0.794*
	5	(20,20,20,20,18)	0.804*	0.801	0.804*
Linear	1	(18,14,10,6,2)	0.531	0.797*	0.705
	1	(18,17,15,13,11)	0.504	0.750*	0.697
	2	(19,18,17,15,13)	0.551	0.756*	0.712
	2	(19,18,14,10,6)	0.567	0.802*	0.718
	3	(19,19,18,17,15)	0.594	0.752*	0.715
	3	(19,19,18,14,10)	0.620	0.804*	0.724
	4	(19,19,19,18,17)	0.643	0.751*	0.719
	4	(19,19,19,18,14)	0.681	0.799*	0.744