TESTING THE EQUALITY OF COEFFICIENT OF VARIATION IN K INVERSE GAUSSIAN POPULATIONS

Amina A. Saleh and Somia K. El-Attar

Department of Applied statistic and Econometric
Institute of Statistical Studies and Research
Cairo University

ABSTRACT

Three tests (likelihood ratio test, Wald test and Score test) are derived for the hypothesis that the coefficients of variation of k Inverse Gaussian populations are equal. The k samples may be of unequal size. The usual χ^2 –approximation is used to investigate the behaviour of the three tests under the null hypothesis and to compare the powers of these tests via a simulation study. Independent samples of equal and unequal size from the Inverse Gaussian distribution were used.

Keywords and Phrases: *Inverse Gaussian distribution; likelihood ratio test; Score test; Wald test; χ^2 -approximation; Coefficient of variation; Simulation of size; Simulation of power.*

1.INTRODUCTION

The coefficient of variation (*CV*) is an important parameter in many physical, biological and medical sciences. In general, it measures the consistency or uniformity of a set of observations on a random variable. Since *CV* is the standard deviation per unit mean, it represents a measure of relative variability. Groups can have the same relative variability even if the means and variances of the variable of interest are different. In several cases more efficient statistical methods can be used, that is when it may be assumed that a number of coefficients of variation of Gaussian (normal) populations are equal but unknown. Naturally then the question arises whether the hypothesis can be tested that the coefficients of variation are equal. Several tested are proposed for the case of normal populations by many authors, for example, Doornbos and Dijkastra (1983), Shafer and Sullivan (1986) and Gupta and Ma (1996).

The inverse Gaussian family denoted IG (μ,λ) is a versatile family for modeling nonnegative right-skewed data, which shares striking similarities with the Gaussian family. For example, analysis of two-factor experiments under an inverse Gaussian model is considered under assumptions equivalent to the equality of variances in the usual normal theory analysis of variance (ANOVA). In this paper we are interested in testing the hypothesis that the coefficients of variation are equal in k inverse Gaussian populations. Such test seems to be missing.

The Inverse Gaussian probability density is given by

$$f(x,\mu,\lambda)=\sqrt{\frac{\lambda}{2\pi}}\,x^{-\frac{3}{2}}\exp[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}],\qquad x>0 \qquad (1.1)$$

where $\mu>0$ and $\lambda>0$.

The square of coefficient of variation (*CV*) for IG(μ_i,λ_i), is given by $\delta_i=\mu_i/\lambda_i$. In section 2, we derive the standard likelihood-ratio test criterion,the Wald test and Rao's Score test for the case of k samples of possibly unequal sizes. Simulation studies are carried out in section 3 for comparing the nominal size and the powers of the various tests described above and the results are displayed in four tables investigating their behaviours under the null hypothesis, using the usual χ^2-approximation as well as comparing the powers of the various tests. The technique of regression analysis as an example of linear model is used to investigate the results of the simulation studies in section 4. Finally, some conclusions are drawn in section 5 and some recommendations are presented. Charts of the powers of the tests shown above are given in the appendix.

2. DEVELOPING TESTS REGARDING THE EQUALITY OF COEFFICIENT OF VARIATION FROM K INVERSE GAUSSIAN POPULATIONS

2.1 THE LIKELIHOOD-RATIO TEST

Suppose that $x_{i1},\ldots,x_{in_i}$; i=1,2, ,k are k independent random samples from X_i~IG(μ_i,λ_i). The likelihood function is

$$L=(2\pi)^{N/2}\left(\prod_{i=1}^{k}\prod_{j=1}^{n_i}x_{ij}^{-3/2}\right)\prod_{i=1}^{k}\left(\frac{\mu_i}{\delta_i}\right)^{n_i/2}\exp[-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{1}{\delta_i\mu_i}\frac{(x_{ij}-\mu_i)^2}{2x_{ij}}]$$

where $N=\sum_i n_i$.

We aim to test the hypothesis H_o: $\delta_i=\delta$; i=1,2,...,k ; for unknown δ, against the alternative H_1: $\delta_i\neq\delta_{i'}$ for at least one pair (i,i').

The log-likelihood under H_o is given by

$$\ln L_o=cons.-\sum_i(n_i/2)\ln\delta+\sum_i(n_i/2)\ln\mu_i-\frac{1}{2\delta}\sum_{i=1}^{k}\frac{n_i}{\mu_i}\left(\bar{x}_i-2\mu_i+\mu_i^2\,\bar{r}_i\right)$$

where $\bar{x}_i=\frac{1}{n_i}\sum_j^{n_i}x_{ij}$ and $\bar{r}_i=\frac{1}{n_i}\sum_j^{n_i}\frac{1}{x_{ij}}$.

Therefore, we have the following likelihood equations:

$$\frac{\partial}{\partial\delta}\ln L_o = \sum n_i - \frac{1}{\tilde{\delta}}\sum_{i=1}^{k}\frac{n_i}{\tilde{\mu}_i}\left(\bar{x}_i - 2\tilde{\mu}_i + \tilde{\mu}_i^2\,\bar{r}_i\right) = 0 \tag{2.1.1}$$

and

$$\frac{\partial}{\partial\mu_i}\ln L_o = \frac{n_i}{\tilde{\mu}_i} - \frac{n_i}{\tilde{\delta}}\left(-\frac{\bar{x}_i}{\mu_i^2} + \bar{r}_i\right) = 0; \qquad i=1,2,\ldots,k. \tag{2.1.2}$$

Simplifying the equations (2.1.1) and (2.1.2), we get

$$\sum\frac{n_i\bar{x}_i}{\tilde{\mu}_i} - \sum n_i = 0 \tag{2.1.3}$$

and

$$\tilde{\mu}_i^{-1} = \frac{\tilde{\delta} + \sqrt{\tilde{\delta}^2 + 4\bar{x}_i\bar{r}_i}}{2\bar{x}_i}; \qquad i=1,2,\ldots,k. \tag{2.1.4}$$

By solving equations (2.1.3) and (2.1.4), we get the restricted maximum likelihood estimates $\tilde{\delta}$ and $\tilde{\mu}_i$'s, of δ and μ_i's respectively, under null hypothesis H_o. Since equation (2.1.3) can not be solved algebraically when k>2, we solve it iteratively by using the Newton's method, with initial value

$$\tilde{\delta}_0 = 2(\bar{x}\bar{r} - 1)$$

where $\bar{x} = \sum_i n_i\bar{x}_i / \sum_i n_i$ and $\bar{r} = \sum_i n_i\bar{r}_i / \sum_i n_i$. This initial value obtained from equation (2.1.3) by replacing $\bar{r}_i$ and $\bar{x}_i$ by $\bar{r}$ and $\bar{x}$ respectively. Our simulation study shows that this initial value is close to the maximum likelihood estimate $\tilde{\delta}$.
Having obtained $\tilde{\delta}$, we use equation (2.1.4) to get $\tilde{\mu}_i$ i=1,2, ... , k.

Therefore, the maximum log-likelihood under H_o is given by

$$\ln\hat{L}_o = (-N/2)\ln(2\pi) + \sum_i\sum_j^{n_i}\ln x_{ij} + \sum_i^{k}\frac{n_i}{2}\ln\left(\tilde{\mu}_i/\tilde{\delta}\right) - \frac{N}{2}$$

To obtain the maximum value of $\ln L$ without the restrictions imposed by H_o, we get first the following unrestricted likelihood equations:

$$\frac{\partial}{\partial\delta_i}\ln L_1 = \frac{1}{\hat{\delta}_i} + \frac{1}{\hat{\delta}_i^2}\left(\frac{\bar{x}_i}{\hat{\mu}_i} - 2 + \bar{r}_i\hat{\mu}_i\right) = 0 \tag{2.1.5}$$

and

$$\frac{\partial}{\partial \mu_i}\ln L_1 = \frac{1}{\hat{\mu}_i} + \frac{\bar{x}_i}{\hat{\delta}_i \hat{\mu}_i^2} - \frac{\bar{r}_i}{\hat{\delta}_i} = 0;\ i=1,2,\ldots,k. \quad (2.1.6)$$

Simplifying equations (2.1.5) and (2.1.6), we get the unrestricted maximum Likelihood estimates $\hat{\delta}_i$ and $\hat{\mu}_i$ of $\hat{\delta}_i$ and μ_i respectively as:

$$\hat{\mu}_i = \bar{x}_i; \qquad \hat{\delta}_i = \bar{x}_i \bar{r}_i - 1 \quad ; \ i=1,2,\ldots,k.$$

Therefore, the maximum log-likelihood without the restrictions imposed by H_o is given by

$$\ln \hat{L}_1 = (-N/2)\ln(2\pi) + \sum_i^k \sum_j^{n_i} \ln x_{ij} + \sum_i^k \frac{n_i}{2}\ln\left(\frac{\bar{x}_i}{\bar{x}_i \bar{r}_i - 1}\right) - \frac{N}{2}$$

Letting λ be the likelihood ratio statistic, we have that

$$-2\ln\lambda = \sum n_i \ln\left(\frac{\bar{x}_i \tilde{\delta}}{(\bar{x}_i \bar{r}_i - 1)\tilde{\mu}_i}\right)$$

It can be shown that $-2\ln\lambda$ is asymptotically distributed as chi-squared with k-1 degrees of freedom (Silvey1970). The decision rules for testing H_o against H_1 at α level of significance would be

$$\text{reject } H_o \text{ if } -2\ln\lambda > \chi^2_{(\alpha;k-1)}$$

2.2. WALD TEST

Let $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$ be an unknown vector-valued parameter and $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_p)$ be the unrestricted M.L.E. of $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$. Suppose the null hypothesis is $H_o' : h(\theta) = [h_1(\theta), h_2(\theta), \ldots, h_p(\theta)] = 0$. Define H as a $p \times m$ matrix with entries $\frac{\partial h_j(\theta)}{\partial \theta_i}, i = 1,2,\ldots,p, j = 1,2,\ldots,m$, and $I(\theta)$ as the fisher information matrix with entries $E\left(-\frac{\partial^2}{\partial\theta_i \partial\theta_j}\ln l\right), i,j = 1,2,\ldots,p$. Then the Wald test statistic is given by $W = h'(\hat{\theta})\left[\hat{H} I^{-1}(\hat{\theta})\hat{H}\right]^{-1} h(\hat{\theta})$, where $h(\hat{\theta}), \hat{H}$ and $I^{-1}(\hat{\theta})$ are the values of $h(\theta), H$ and $I^{-1}(\theta)$ when θ is replaced by $\hat{\theta}$, see Wald (1943). Furthermore, W is asymptotically distributed as χ^2 with m degrees of freedom (Silvey, page 116).

In our case, $\theta=(\mu_1,\lambda_1,\mu_2,\lambda_2,...,\mu_k,\lambda_k)$ and $h_i(\theta)=\frac{\mu_{i+1}}{\lambda_{i+1}}-\frac{\mu_i}{\lambda_i}$, $i=1,2,...,k-1$.
Under H_o', which is equivalent to H_o, W is asymptotically distributed as χ^2 with k-1 degrees of freedom. When k=2,

$$W_2=\frac{(\hat{\delta}_1-\hat{\delta}_2)^2}{a_1+a_2},$$

When k=3,

$$W_3=\frac{a_1(\hat{\delta}_2-\hat{\delta}_3)^2+a_2(\hat{\delta}_1-\hat{\delta}_3)^2+a_3(\hat{\delta}_1-\hat{\delta}_2)^2}{a_1a_2+a_1a_3+a_2a_3},$$

When k=4,

$$W_4=\frac{1}{a_1a_2a_3+a_1a_2a_4+a_1a_3a_4+a_2a_3a_4}\Big[a_1a_2(\hat{\delta}_3-\hat{\delta}_4)^2+a_1a_3(\hat{\delta}_2-\hat{\delta}_4)^2$$

$$+a_1a_4(\hat{\delta}_2-\hat{\delta}_3)+a_2a_3(\hat{\delta}_1-\hat{\delta}_4)^2+a_2a_4(\hat{\delta}_1-\hat{\delta}_3)^2+a_3a_4(\hat{\delta}_1-\hat{\delta}_2)^2\Big]$$

where $\hat{\delta}_i=\bar{x}_i\bar{r}_i-1$ and $a_i=\frac{\delta_i^3}{n_i}+\frac{2\delta_i^2}{n_i}$, $i=1,2,...,k$. A generalized formula for W_k can be easily deduced from the above formulas.

2.3. RAO'S SCORE TEST

Let $\theta=(\delta_1,\delta_2,...,\delta_k,\mu_1,\mu_2,...,\mu_k)$, $U(\theta)=\left(\frac{\partial}{\partial\delta_1}\ln L,\frac{\partial}{\partial\delta_2}\ln L,...,\frac{\partial}{\partial\delta_k}\ln L,\right.$ $\left.\frac{\partial}{\partial\mu_1}\ln L,\frac{\partial}{\partial\mu_2}\ln L,....,\frac{\partial}{\partial\mu_k}\ln L,\right)$, $B(\theta)$ be the $2k\times2k$ matrix with entries $\left(-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln L\right)$ and $I(\theta)=E(B(\theta))$. Under the null hypothesis $H_o:\delta_1=\delta_2=...=\delta_k$, we get the restricted M.L.E.of θ i.e. $\tilde{\theta}=(\tilde{\delta}_1,\tilde{\delta}_2,...,\tilde{\delta}_k,\tilde{\mu}_1,\tilde{\mu}_2,...,\tilde{\mu}_k)$ by solving the equations (2.1.3) and (2.1.4). The Rao's Score test statistic is given by $R_{sc}=U'(\tilde{\theta})I^{-1}(\tilde{\theta})U(\tilde{\theta})$. Also, R_{sc} is asymptotically distributed as χ^2 with k-1 degrees of freedom, see Rao(1973) and Lawless(1982).

In our case, the scores are

$$\frac{\partial}{\partial\delta_i}\ln L_1=\frac{-n_i}{2\delta_i}+\frac{n_i}{2\delta_{,i}^2}\left(\frac{\bar{x}_i}{\mu_i}-2+\bar{r}_i\mu_i\right), i=1,2,...,k$$

and

$$\frac{\partial}{\partial \mu_i}\ln L_1 = \frac{n_i}{2\mu_i} + \frac{n_i \bar{x}_i}{2\delta_i \mu_i^2} - \frac{n_i \bar{r}_i}{2\delta_i}, i = 1,2,\ldots,k$$

The Fisher information matrix $I(\theta)$ is

$$I(\theta) = \begin{pmatrix} I_{11}(\theta) & I_{12}(\theta) \\ I_{21}(\theta) & I_{22}(\theta) \end{pmatrix}$$

where $I_{11}(\theta) = diag\left(\frac{n_i}{2\delta_i^2}\right)$, $I_{12}(\theta) = I_{21}(\theta) = diag\left(\frac{-n_i}{2\delta_i \mu_i}\right)$, and $I_{22}(\theta) = diag\left(\frac{n_i(\delta_i + 2)}{2\delta_i \mu_i^2}\right)$.

Let $b_i = \frac{-n_i}{2\tilde{\delta}_i} + \frac{n_i}{2\tilde{\delta}_i^2}\left(\frac{\bar{x}_i}{\tilde{\mu}_i} - 2 + \bar{r}_i \tilde{\mu}_i\right), \quad i = 1,2,\ldots,k$.

and

$$c_i = \frac{n_i}{2\tilde{\mu}_i} + \frac{n_i \bar{x}_i}{2\tilde{\delta}_i \tilde{\mu}_i^2} - \frac{n_i \bar{r}_i}{2\tilde{\delta}_i} = 0, \quad i = 1,2,\ldots,k$$

Thus $U(\tilde{\theta}) = (b_1, b_2, \ldots, b_k, \ldots, c_1, c_2, \ldots, c_k)$. The explicit form of the test statistic is given by:

$$R_{sc} = \tilde{\delta}^2(\tilde{\delta} + 2)\sum_i \frac{b_i^2}{n_i}$$

Which is approximately distributed as $\chi^2_{(k-1)}$.

3.SIMULATION RESULTS

A simulation study was used to compare the estimated size of the test, using the χ^2-distribution as an approximation, to the nominal size and to obtain estimates of the power for various alternatives. In this simulation study, samples were taken from Inverse Gaussian distributions using the method given by Michael, et.al.(1976). We restricted ourselves to the cases k=2 and k=4. Various sample size patterns were used. Balanced patterns used sample sizes of "5,10,15, 20,30" for both cases. The unbalanced patterns used samples of sizes "5,25" and "25,5" for the case k=2 and " 5,10,15,25" and "25,15,10,5" for the case k=4.

Table I shows the simulation results for the estimated size of the likelihood ratio test for $H_o: \delta_i = \delta$; i=1,2 , for seven values of δ . Table II gives similar results for the case of four samples. These results are based upon one thousand replications. The nominal size was set at .05.

Table I

Estimated Size Using Two Samples from Inverse Gaussian Distributions
Nominal Size =.05 with 1000 replications

CV^2		.1	.5	1	2	3	5	10
$n_i = 5$	LR	.099	.095	.088	.088	.089	.096	.099
	ST	.023	.021	.014	.007	.007	.003	.003
	WT	0	0	0	0	0	0	0
10	LR	.090	.057	.067	.061	.073	.063	.065
	ST	.066	.050	.048	.028	.018	.006	.003
	WT	0	0	0	0	0	0	0
20	LR	.050	.053	.048	.052	.051	.053	.048
	ST	.040	.041	.051	.029	.023	.016	.005
	WT	.019	.008	.003	0	0	0	0
25	LR	.025	.053	.044	.050	.048	.052	.047
	ST	.050	.051	.050	.042	.037	.021	.008
	WT	.027	.019	.007	0	0	0	0
30	LR	.053	.057	.050	.047	.069	.057	.069
	ST	.051	.055	.046	.049	.044	.027	.006
	WT	.031	.028	.011	.002	0	0	0
5,25	LR	.092	.109	.078	.090	.084	.105	.110
	ST	.026	.019	.022	.037	.055	.121	.226
	WT	.023	.002	0	0	0	0	0
25,5	LR	.107	107	.121	.084	.115	.084	.098
	ST	.038	.019	.022	.030	.059	.091	.207
	WT	.02	.014	.005	.002	0	0	0

It can be seen from this table that the estimated size for the LR test is larger than the nominal for small samples size (5 and 10 for the balanced case and "5, 25" and "25, 5" for the unbalanced case). As the sample size get bigger, the results for this test are improving quickly for the balanced case. For most values of the coefficient of variation, the best results are obtained at sample sizes 20, 25 and 30. The estimated size for the ST is smaller than the nominal for small sample size (5 and 10 for the balanced case and "5, 25" and "25,5" for the unbalanced case), On the other direction, for equal sample sizes, the estimated size is decreasing as the value of CV increasing.

For unequal sample sizes, it is decreasing as the value of *CV* approach the value 1, but for the values of *CV* greater than 1, it is increasing with the increasing value of *CV*. The best results are obtained for sample sizes 20, 25 and 30 and *CV* less than 1 The *WT* tends to be too conservative, i.e. the actual size is (much) smaller than the nominal size.

Table II

Estimated Size Using Four Samples from Inverse Gaussian Distributions
Nominal Size =.05 with 1000 replications

CV^2		.1	.5	1	2	3	5	10
$n_i = 5$	LR	.150	.110	.136	.148	.138	.134	.153
	ST	.056	.050	.052	.046	.043	.023	.013
	WT	0	0	0	0	0	0	0
10	LR	.074	.088	.076	.089	.085	.088	.098
	ST	.048	.054	.056	.050	.040	.025	.018
	WT	.036	.012	.002	.002	0	0	0
20	LR	.052	.048	.050	.052	.060	.050	.054
	ST	.038	.052	.041	.048	.028	.031	.016
	WT	.01	.012	.014	.004	.001	0	0
25	LR	.072	.048	.056	.052	.052	.050	.052
	ST	.050	.052	.051	.048	.050	.038	.020
	WT	.05	.024	.022	.016	.002	.001	0
30	LR	.076	.064	.056	.063	.059	.056	.052
	ST	.078	.052	.050	.051	.049	.050	.052
	WT	.052	.042	.021	.022	.006	.002	0
5,10, 15,25	LR	.104	.097	.101	.078	.086	.073	.083
	ST	.043	.038	.043	.099	.174	.363	.511
	WT	.173	.162	.127	.093	.090	.030	.013
25,15, 10,5	LR	.095	.083	.085	.081	.085	.080	.075
	ST	.044	.034	.040	.110	.201	.341	.477
	WT	.187	.158	.140	.106	.090	.032	.012

Table II demonstrate the estimated size for the three testes for the case of four samples. For the *LR* test the estimated size is larger than the nominal for most cases. The best results are obtained for the balanced case with sample sizes 20, 25 and 30 where the estimated size approached the nominal size as the value of *CV* get bigger. The estimated size for the *ST* is closer to the nominal size when n is small (5, 10) and the *CV* less than 3 In general, the *ST* is somewhat better than the *LR* test. For unbalanced case, the estimated size is increasing, as the value of *CV* gets bigger. The best results are obtained for sample sizes 25 and 30 with small values of *CV*. The *WT* give estimated size closer to the nominal size for n=25 and 30 with small value of *CV* (1), in the other cases it is too conservative.

Tables III and IV include the estimated power for two and four samples respectively, and different combinations of *CV*'s for the alternative hypotheses. These simulations also used one thousand replications and a .05 significance level.

Here we estimated the probability of rejected the hypothesis for several patterns of *CV's*. The sample sizes are the same as in the previous simulations studies, but we add n =50, 100, for the balanced case.

Table III

Estimated Power Using Two Samples from Inverse Gaussian Distributions
Nominal Size =.05 with 1000 replications

CV^2		a_1	a_2	A_3	a_4	a_5	a_6	a_7
$n_i = 5$	LR	.399	.212	.143	.128	.212	.197	.409
	ST	.161	.250	.020	.016	.013	.012	.008
	WT	0	0	0	0	0	0	0
10	LR	.635	.314	.177	.160	.263	.459	.804
	ST	.556	.236	.118	.082	.086	.132	.082
	WT	.003	0	0	0	0	0	0
20	LR	.908	.565	.276	.187	.429	.755	.966
	ST	.887	.526	.229	.142	.334	.638	.649
	WT	.764	.189	.022	0	0	0	0
25	LR	.946	.690	.326	.263	.526	.769	.992
	ST	.941	.658	.303	.229	.459	.690	.800
	WT	.898	.458	.116	.008	.015	.006	.001
30	LR	.973	.711	.357	.289	.556	.821	.997
	ST	.972	.689	.332	.257	.520	.790	.902
	WT	.961	.542	.179	.067	.060	.045	.001
50	LR	.998	.928	.542	.494	.780	.956	1
	ST	.998	.910	.496	.542	.836	.972	1
	WT	.996	.898	.444	.310	.558	.718	.154
100	LR	1	1	.916	.834	.994	1	1
	ST	1	1	.904	.882	.996	1	1
	WT	1	1	.902	.768	.988	1	.998
5,25	LR	.651	.220	.121	.183	.302	.372	.559
	ST	0	.303	.160	.002	.001	.001	.001
	WT	.881	.033	.079	.233	.229	.104	.018
25,5	LR	.399	.366	.217	.107	.178	.313	.554
	ST	.530	.001	.001	.110	.227	.412	.655
	WT	.025	.572	.367	.071	.063	.027	.013

a_1: .1,.5 a_2: 1,.3 a_3: 1,.5 a_4: 1,2 a_5: 1,3 a_6: 1,5 a_7: 1,10

These tables suggested the following conclusion:

1. For very small samples (i.e. n=5,10) WT has no power at all.
2. For increasing sample size, the results for these tests are improving quickly.
3. For large samples (25 or more), the differences in power between these tests are minimal.
4. The power becomes larger for those alternatives that represent more separation of the *CV's*, while it is smaller when the *CV's* are closer to each other.

The charts of the power are available in the appendix.

Table IV

Estimated Power Using Four Samples from Inverse Gaussian Distributions
Nominal Size =.05 with 1000 replications

CV^2		A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
$n_i = 5$	***LR***	.254	.215	.202	.246	.258	.212	.242	.332	.360	.372	.343	.313	.606	.532	.452
	ST	.148	.089	.068	.110	.114	.086	.058	.076	.056	.142	.228	.071	.112	.326	.030
	WT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	LR	.338	.28	.206	.302	.288	.226	.342	.414	.602	.600	.535	.408	.920	.874	.644
	ST	.24	.146	.126	.196	.267	.122	.170	.166	.168	.312	.530	.106	.314	.866	.074
	WT	.106	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	LR	.736	.411	.264	.55	.510	.312	.572	.710	.870	.872	.783	.654	.998	.980	.942
	ST	.610	.305	.204	.436	.548	.158	.352	.364	.360	.770	.813	.202	.900	.986	.152
	WT	.598	.022	.026	.032	.039	.028	.034	.006	0	.020	.012	.008	.006	.002	0
25	LR	.788	.555	.332	.634	.550	.432	.654	.864	.948	.952	.843	.769	.974	.998	.978
	ST	.704	.394	.254	.574	.574	.236	.398	.556	.540	.912	.863	.323	.974	.999	.286
	WT	.706	.116	.070	.068	.054	.134	.182	.142	.066	.071	.032	.172	.104	.030	.017
30	LR	.920	.655	.432	.652	.642	.506	.766	.940	.996	.976	.919	.890	.999	.998	.997
	ST	.860	.517	.354	.598	.642	.314	.514	.706	.694	.960	.917	.414	.982	.998	.293
	WT	.874	.240	.190	.190	.102	.266	.326	.396	.298	.240	.088	.456	.026	.036	.113
50	LR	.990	.870	.678	.930	.852	.786	.940	.998	1	.994	.980	.994	1	1	1
	ST	.974	.794	.566	.928	.862	.590	.822	.960	.966	.996	.978	.848	1	1	.682
	WT	.984	.696	.508	.718	.292	.716	.830	.910	.932	.896	.262	.956	.390	.132	.956
100	LR	1	.998	.966	.998	.986	.994	1	1	1	1	1	1	1	1	1
	ST	.998	.992	.954	.998	.988	.974	1	1	1	1	1	1	1	1	1
	WT	1	.996	.948	.998	.898	.992	1	1	1	1	.992	1	1	.380	1
5,10, 15,25	LR	.434	.276	.236	.354	.516	.172	.276	.386	.390	.792	.764	.254	.876	.976	.356
	ST	.134	.026	.034	.050	.260	.032	.014	.018	.010	.272	.434	.048	.030	.390	.016
	WT	.592	.156	.196	.1660	.164	.142	.116	.072	.024	.164	.110	.106	.012	.050	.022
25,15, 10,5	LR	.510	.282	.218	.262	.162	.328	.432	.570	.790	.618	.306	.696	.932	.470	.961
	ST	.538	.314	.224	.330	.232	.310	.404	.538	.674	.676	.396	.460	.950	.564	.522
	WT	.088	.044	.060	.034	.048	.044	.016	.018	.014	.028	.036	.028	.016	.016	.006

A_1: .1,.2,.3,.4 A_2: 1,2,3,4 A_3: 1,2,2,3 A_4: 1,1,3,3 A_5: 1,1,1,3
A_6: 1,3,3,3 A_7: 1,3,3,5 A_8: 1,3,5,7 A_9: 1,4,7,10 A_{10}: 1,1,5,5
A_{11}: 1,1,1,5 A_{12}: 1,5,5,5 A_{13}: 1,1,10,10 A_{14}: 1,1,1,10 A_{15}: 1,10,10,10

4.MORE STATISTICAL ANALYSIS FOR POWER AND SIZE

Using the simulations results, we tried to use the regression analysis as an example of linear models to obtain more information about the power and size of the above three tests

4.1.Regression analysis for the size:

We regress the size of the above tests on the following variables: the size of the sample, n; number of samples, k ; coefficient of variation, CV ; the parameter μ; and the parameter, λ

Using the stepwise procedure, we found the best models for the LR and ST are that ones which contains n, k and CV with fitted models:

$$Size(LR) = .09418 - .002135\ n + .004563\ k + .0008287\ CV$$

$$Size(ST) = .01132 + .0006114\ n + .007456\ k - .003796\ CV$$

For the WT the fitted model is :

$$Size(WT) = -01042 + .000459\ n + .002789\ k + .0004761\ CV + .001442\ \lambda - .001321\ \mu$$

Where all independent variables are very significant (p-value = zero except for μ for the last model it is equal .024). Also by performing multiple regression we obtained the same conclusions.

4.2.Regression analysis for the power:

For each cases, k=2 and k=4, we regress the power of the above tests on the following variables: the size of the sample, n; and the coefficient of variations, CV"s ; Using the stepwise procedure we found that , for the LR and ST, that all variables are significant . The fitted models are:

For k = 2:

$$Power(LR) = .637 + .0064\ n - .41\ CV_1 + .04154\ CV_2$$

$$Power(ST) = .515 + .008705\ n - .397\ CV_1 + .02207\ CV_2$$

While for the WT the fitted model is :

$$Power(WT) = -.359 + .02537\ n$$

For K = 4:

$$Power(LR) = .521 + .0061\ n - .328\ CV_1 + .05178\ CV_4$$

$$Power(ST) = .37 + .008564\ n - .2261\ CV_1 - .0595\ CV_2 + .04758\ CV_4$$

While for the WT the fitted model is :

$$Power(WT) = .274 + .01095\ n - .429\ CV_1 + .01822\ CV_2$$

Where all independent variables are very significant.

5-CONCLUSON

We conclude that LR test has the highest power among the other two. The ST is only second to the LR test. The WT's power decreases sharply when the sample size decreases. For large sample sizes (25 or more) LR and ST give the user excellent control over the size and their power is quite satisfactory .

BIBLIOGRAPHY

Doornbos, R. and Dijkstra,J.B. (1983)," A Multi Sample Test for the Equality of Coefficient of Variation in Normal Populations" *Communication in Statistics: Simulation and Computation*, Vol.12, No2,147-158.

Gupta, Ramesh C. and Ma, Suchun (1996),"Testing the Equality of Coefficient of Variation in k Normal Populations", *Communication in Statistics : Theory and Methods.* Vol.25 No.1, (115-132).

Lawless,J.K. (1982), *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons,Inc..

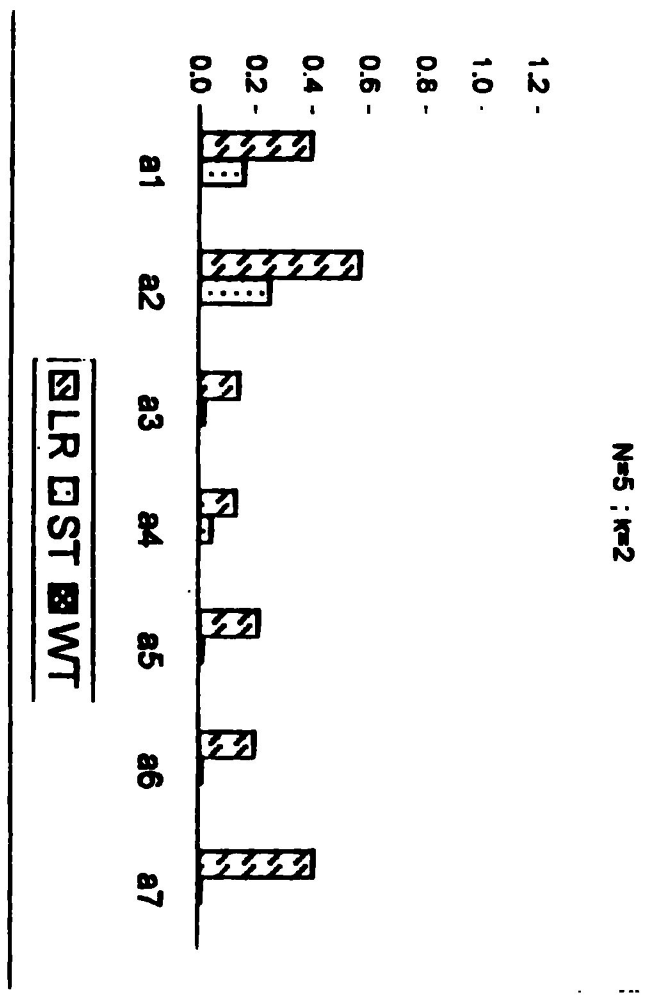
Michael, John R. , Schucany, William R. and Hass, Roy W. (1976), "Generating Random Variates Using Transformations With Multiple Roots", *The American Statistician*, Vol.30,No.2 (88- 90).

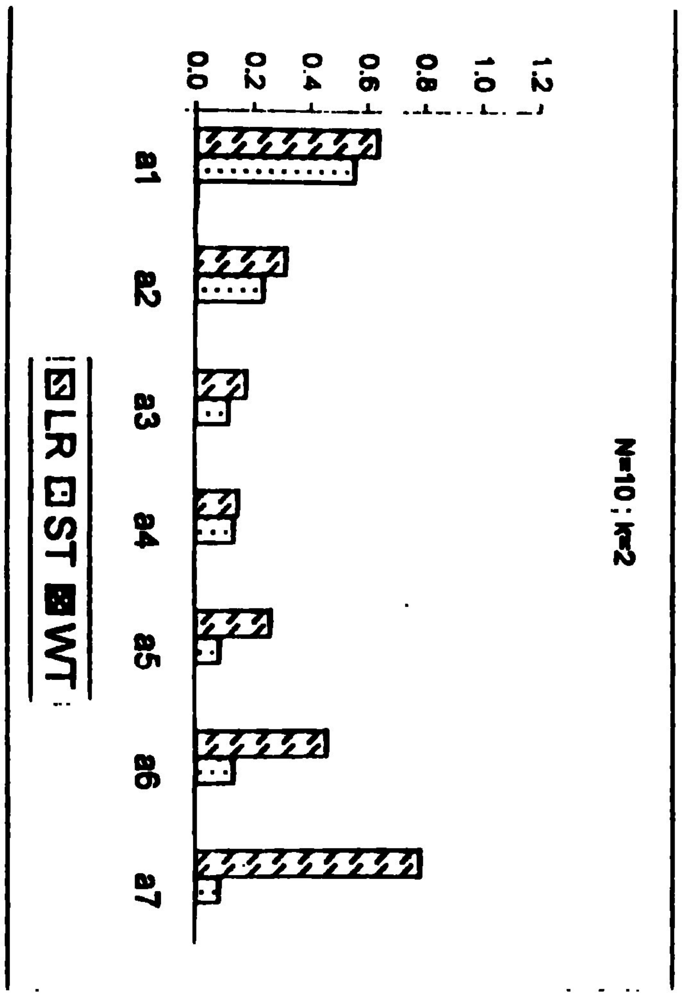
Rao,C.R.(1973), *Linear Statistical Inference and Its Applications*, John Wiley & Sons,Inc..

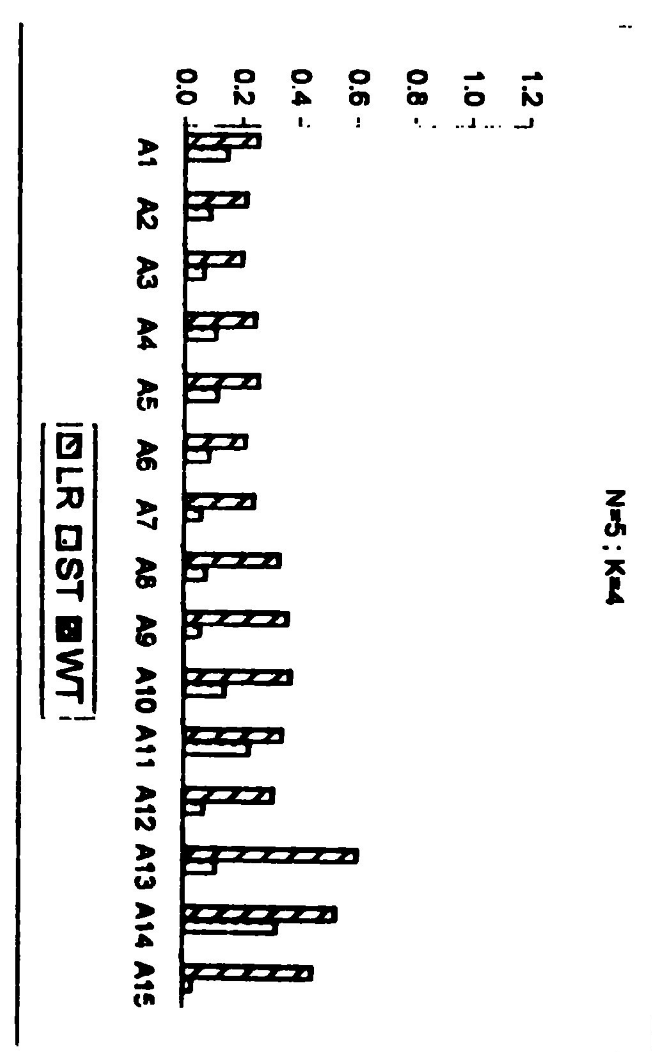
Shafer, N.J and Sullivan,J.A. (1986), " A Simulation Study of a Test for the Equality of the Coefficients of Variation ", *Communication in Statistics: Simulation and Computation*,. Vol.15, No3,681-695.

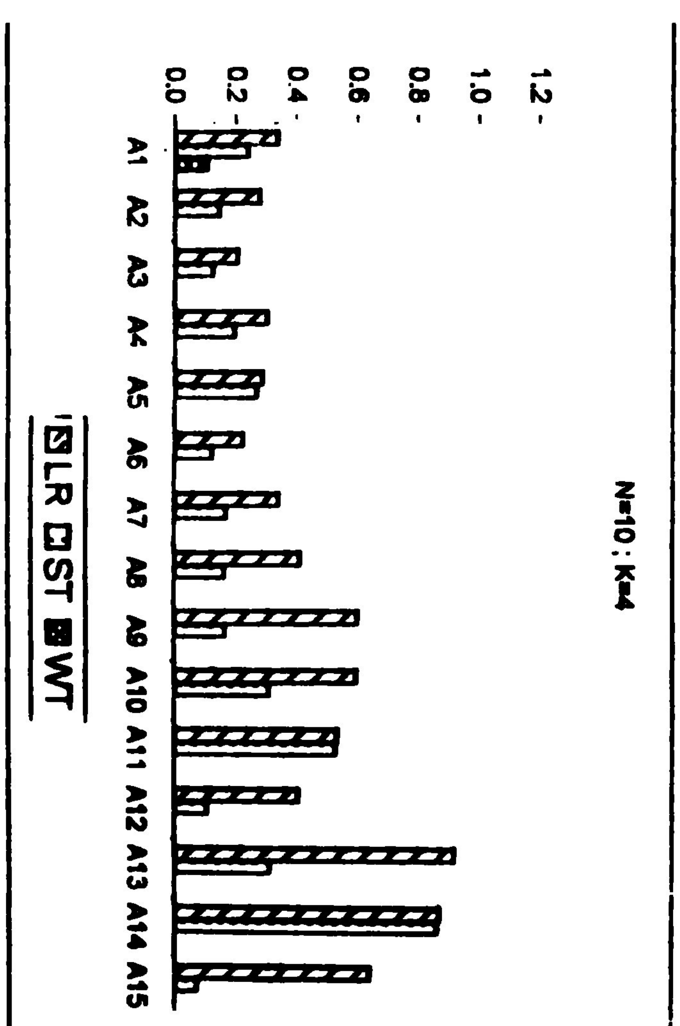
Silvey , S.D. (1970), *Statistical Inference*, Chapman and Hall.

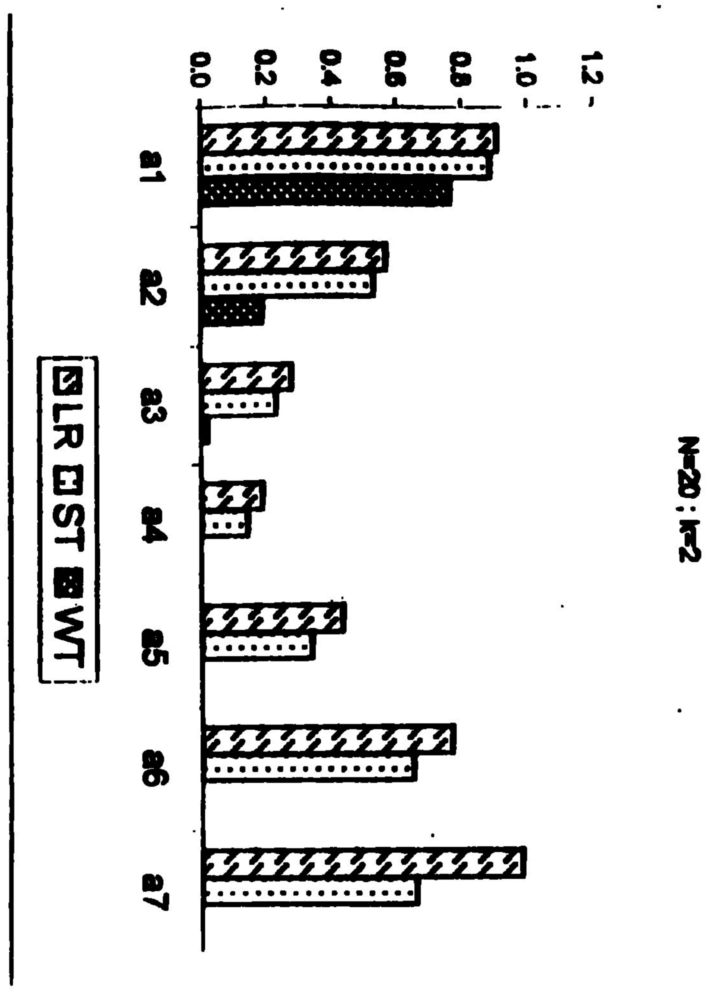
Wald,A.(1943), " Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is large", *Trans. Am. Math. Soc.*, Vol . 54, 426-482.

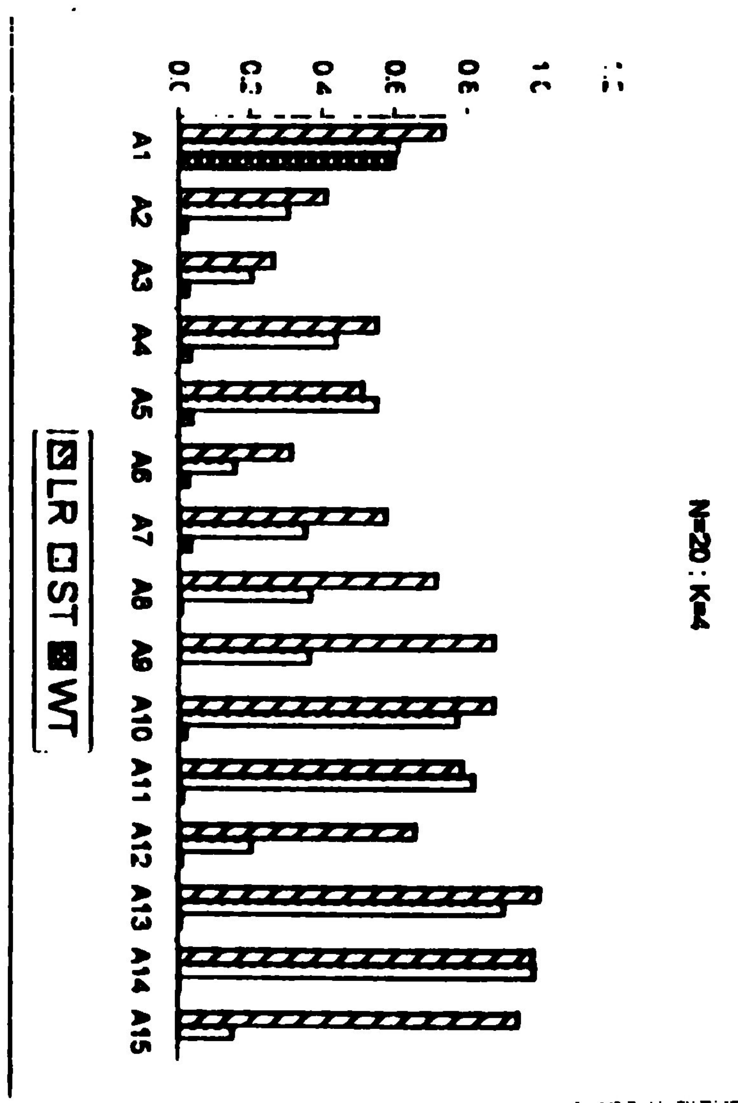

N=5 ; k=2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
a1 a2 a3 a4 a5 a6 a7
LR ST WT

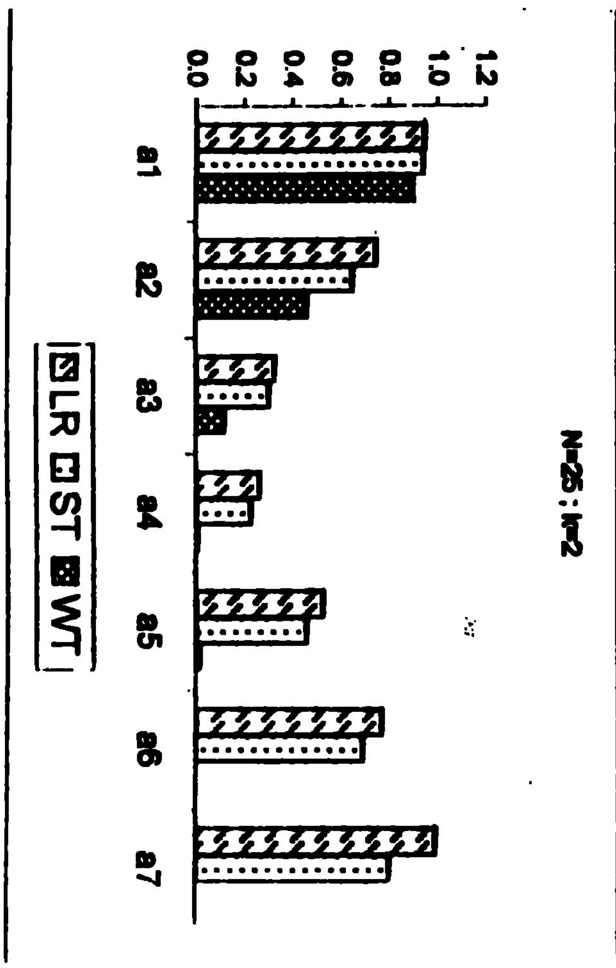

N=10 ; k=2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
a1 a2 a3 a4 a5 a6 a7
LR ST WT

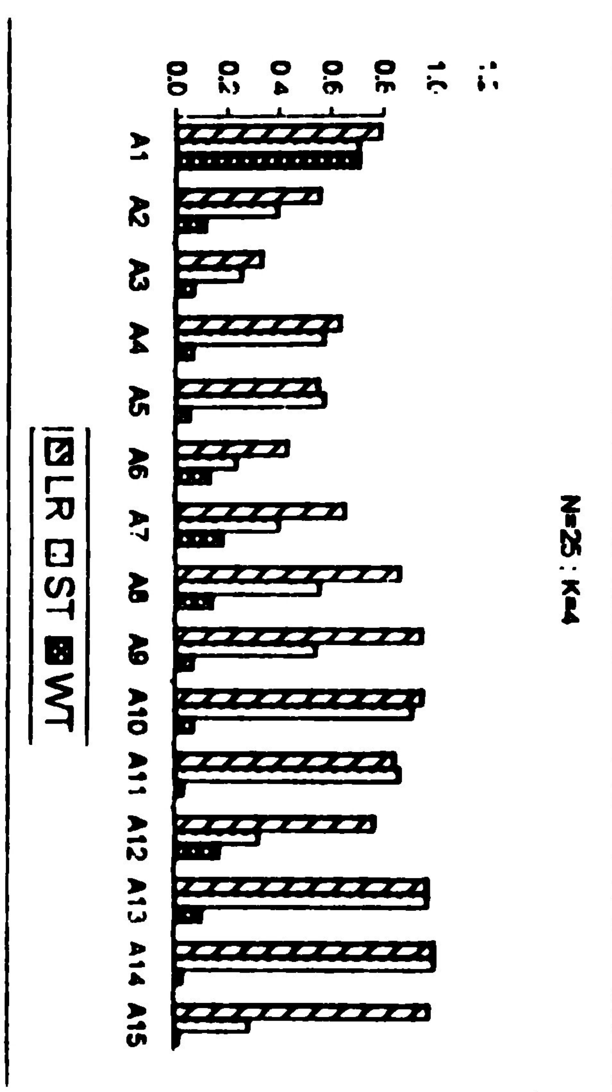

N=5 ; K=4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15
LR ST WT

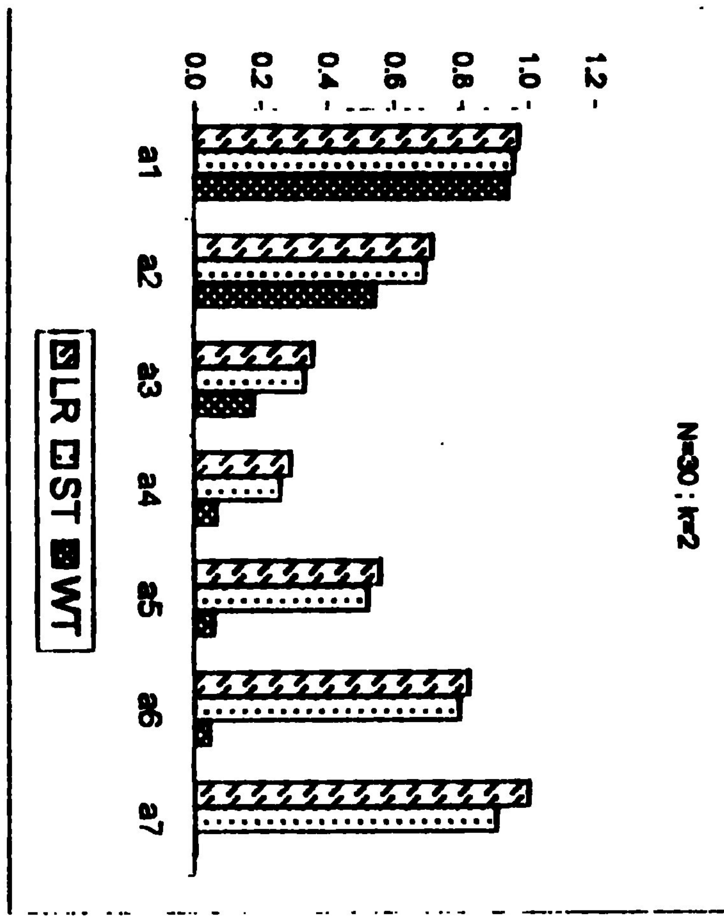

N=10 ; K=4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15
LR ST WT

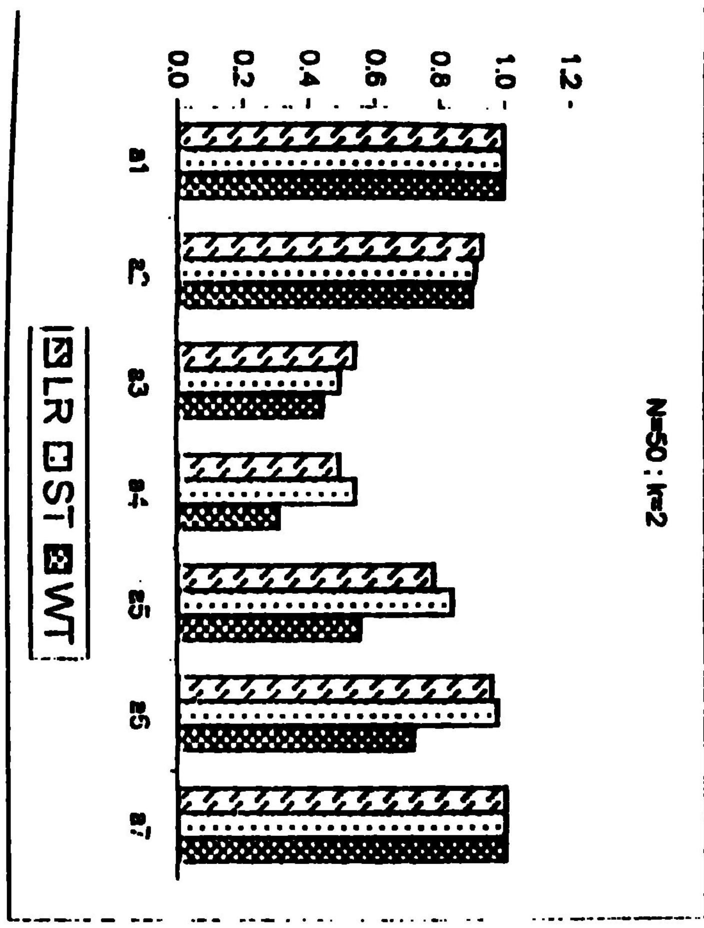

N=20 ; k=2
LR ST WT

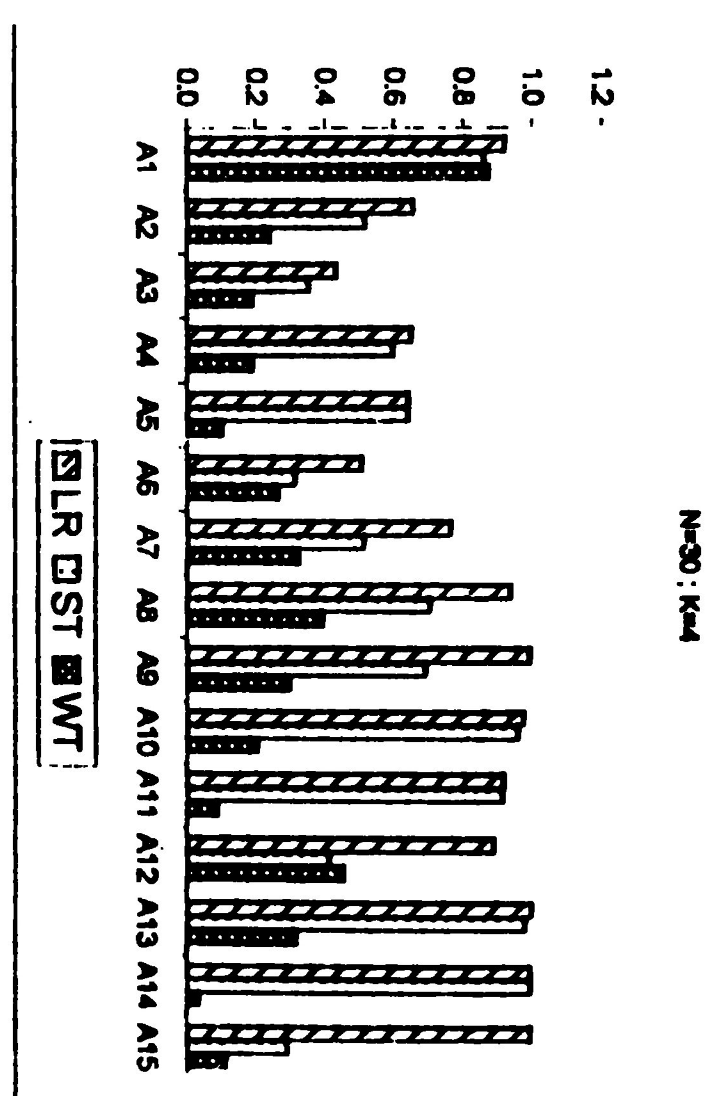

N=20 : K=4
LR ST WT

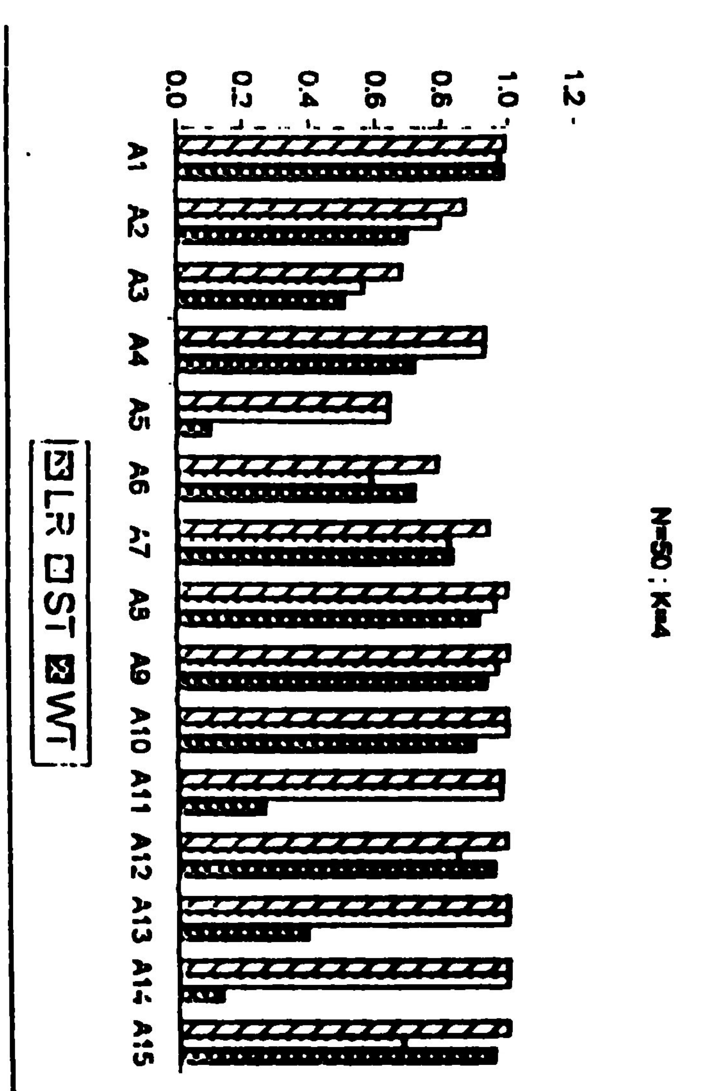

N=25 ; k=2
LR ST WT

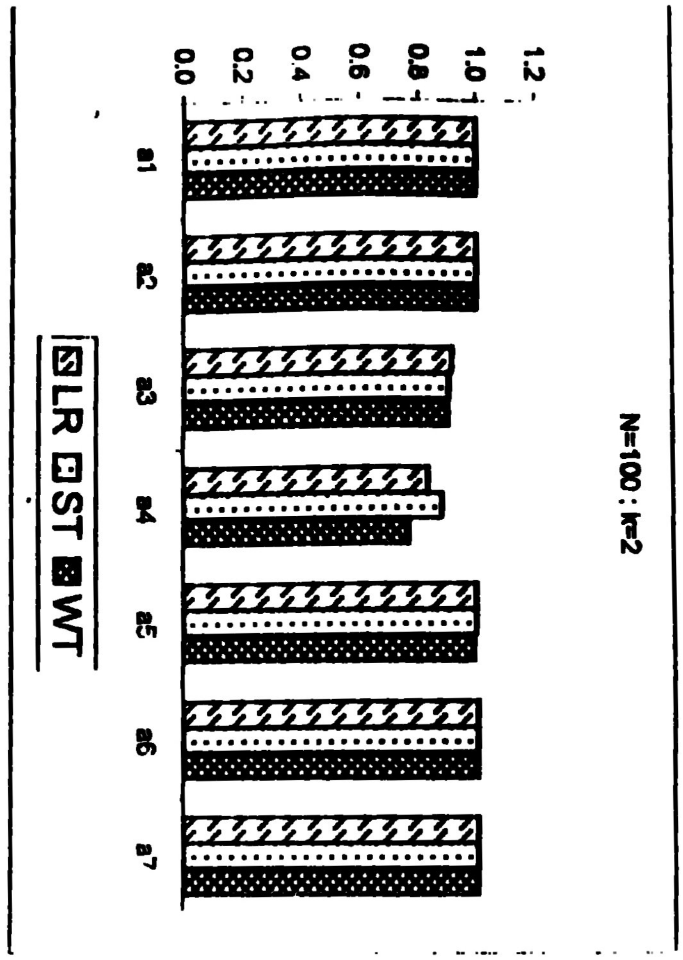

N=25 : K=4
LR ST WT

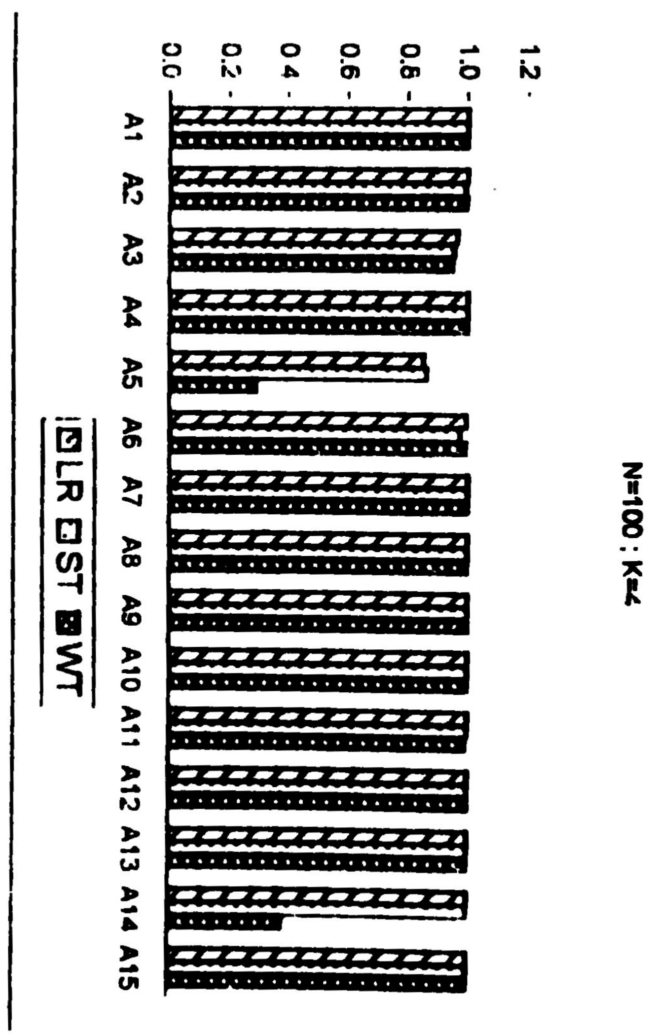

N=30 ; k=2
1.2
1.0
0.8
0.6
0.4
0.2
0.0
a1 a2 a3 a4 a5 a6 a7
LR ST WT


N=50 ; k=2
1.2
1.0
0.8
0.6
0.4
0.2
0.0
LR ST WT


N=30 ; K=4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15
LR ST WT


N=50 ; K=4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A1 A2 A3 A4 A5 A6
LR ST WT


N=100 : k=2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
a1
a2
a3
a4
a5
a6
a7
LR ST WT


N=100 : K=4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
A11
A12
A13
A14
A15
LR ST WT