TESTING FOR SERIAL CORRELATION BY MEANS OF EXTREME VALUES

Ishay Weissman

•

Faculty of Industrial Engineering and Management
Technion - Israel Institute of Technology

e-mail: ieriw01@ie.technion.ac.il

Dedicated to the hundredth birthday of Boris Gnedenko

ABSTRACT

The *largest spacing* of a sample is suggested as a possible test-statistic to detect serial dependence (correlation) among the data. A possible application is in testing the quality of random number generators, which are so important in the study of systems reliability. We compare its performance to the *Kolmogorov-Smirnov* test because of their similar nature – one is based on extreme distance between order statistics, the second on extreme discrepancy between the empirical distribution function and the theoretical one. The tests are applied to several models with serial dependence. Special attention is given to an *autoregressive model*. Based on Monte Carlo simulations, the largest spacing is more powerful for moderately large sample size, over 50, say. A surprising connection to extreme values is discovered, namely, that the *likelihood-ratio* test, which is most powerful under the autoregressive alternative, is based on lower extremes.

Keywords: Autoregressive model, Binomial model, Kolmogorov-Smirnov test, Largest spacing, Likelihood-ratio test, Monte Carlo methods, Moving-max model.

1   INTRODUCTION AND MOTIVATION

Suppose we have a sample $X_1, X_2, \ldots, X_n$ from some continuous distribution function $F0$. Suppose further that $F0$ is the uniform distribution over the unit interval $[0, 1]$ (if not, we replace the Xi by $F_0(X_i)$). Under ideal conditions we expect that the sample is an iid sample, but we suspect that the data at hand exhibit some serial correlation among consecutive observations and we want to put it under a statistical test.

So, let $H0$ denote the null hypothesis of "iid-uniform". If the alternative is "not iid-uniform", then, as J.E. Gentle says, "*this alternative is uncountably composite and there cannot be a most powerful test. We need a suite of statistical tests. Even so, of course, not all alternatives can be addressed*" (Gentle (2003), page 71). In the context of random number generation L'Ecuyer & Simard (2007) offer a comprehensive battery of tests called **TestU01**. The authors state on page 4 that "*the number of different tests that can be defined is infinite and these tests detect different problems.*" Our aim in this paper is to attack a narrower problem, namely the existence of serial correlation. There are many possible models which possess serial correlation. To be specific, we start with the *autoregressive model*

$$X_i = \rho X_{i-1} + (1 - \rho)U_i \quad (1 \le i \le n, \, 0 \le \rho < 1) , \tag{1}$$

where $\{ U_i : i \ge 0 \}$ is a $U[0,1]$-iid sequence and $X_0 = U_0$. In this setting, the problem is to test

$$H_0 : \rho = 0 \ \text{ vs. } \ H_1 : \rho > 0 . \tag{2}$$

In Section 6 we deal with the *binomial model* and the *moving-max model*.

Back to Equation (1), a plot of consecutive pairs, X_{i+1} vs. X_i, is very useful here. For instance, on the left plot of Figure 1, we have $n = 2000$ pairs with $\rho = 0$ and on the right, the same with $\rho = 0.1$. In the latter, the slopes below and above the data are quite visible. But as ρ decreases to 0, detection of existence of a serial correlation by the human eye becomes more and more difficult.

n=2000, rho=0.0   **n=2000, rho=0.1**
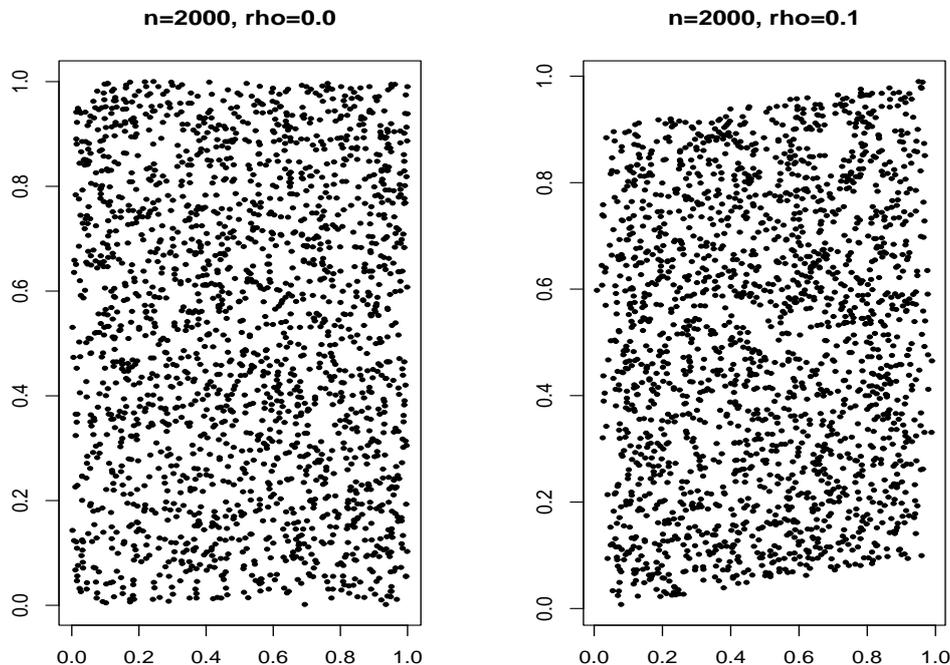
Figure 1. Pairs of Successive Numbers, X_{i+1} vs. X_i

This paper evolved as a result of the author's empirical observation that the largest spacing (LS) of a sample (see a formal definition in Section 2) is quite effective in detecting serial correlation in the sample (see Onn and Weissman (2011)). In order to study its performance with respect to other test statistics, we compare it to the Kolmogorov-Smirnov (KS) test. The KS test is chosen not only because it is so widely used as a goodness of fit test, but also because of its similar nature to LS. That is, while KS distance is the extreme vertical distance between the empirical distribution and the uniform distribution, LS is the extreme horizontal distance between consecutive order statistics. We also compare the two tests to the performance of the sample serial correlation (SSC), the least squares estimator of ρ. Under normality the latter is the natural one to use and it would be interesting to compare the performance of the first two tests to the performance of SSC under the present set-up. The comparisons in terms of power are presented in Section 3.

We then ask what is a most powerful test for the null hypothesis (2). It is somehow surprising that the answer is again an extreme statistics. In Section 4 we define the transformed data $\{T_{1i} : 1 \leq i \leq n\}$, which are iid, and show that the likelihood-ratio test (LRT) (which is most powerful) is based on $\min_{1 \leq i \leq n} T_{1i}$. In Section 5 we discuss similar issues under the *autoregressive model of order k*. Two more models which exhibit serial dependence are introduced in Section 6 and the performance of the LS and KS tests applied to these models are discussed.

Remark 1. The $\{X_i\}$ as defined by (1) are marginally not $U[0, 1]$-distributed if $\rho > 0$. But if the $\{U_i\}$ are discrete uniform (rather than continuous uniform), the $\{X_i\}$ can be uniform too: Suppose Y_1 and Y_2 are both $U[0, 1]$-distributed and the error term ε is a random variable, independent of Y_1. Suppose further that for some $\rho > 0$ one has

$$Y_2 = \rho Y_1 + (1 - \rho)\varepsilon.$$

What are the conditions on ρ and ε for this to hold? Lawrance (1992) proves that $\rho = k^{-1}$ for some integer $k \geq 2$ and ε must be uniform over the set $\{0, 1, \dots, k-1\}/(k-1)$. Hence, for very large k, ε is also (approximately) $U[0, 1]$-distributed. In fact, computer generated "random numbers" are of ε-type.

2 ASYMPTOTIC BACKGROUND

Largest spacing. Let $Y_1 \leq Y_2 \leq \dots \leq Y_n$ be the order statistics of $\{X_1, X_2, \dots, X_n\}$ and Y_0, Y_{n+1} be given by 0, 1 respectively. Define the *sample spacings* $V_i = Y_i - Y_{i-1}$ ($i = 1, 2, \dots, n+1$). Darling (1953) studies properties of the sample spacings (under H_0) in the context of a random partition of an interval. In particular he gives the exact joint distribution of (V_{min}, V_{max}),

$$P\{V_{\min} \geq x, \ V_{\max} \leq y\} = \sum_{j=0}^{n+1} (-1)^j \binom{n+1}{j} \{(1 - (n+1-j)x - jy)_+\}^n.$$

Putting $x = 0$ we obtain the Whitworth (1897) result

$$P\{V_{\max} \leq y\} = \sum_{j=0}^{n+1} (-1)^j \binom{n+1}{j} \{(1 - jy)_+\}^n. \tag{3}$$

and putting $y = 1$, we get

$$P\{V_{\min} \geq x \} = \{1 - (n+1)x\}^n \qquad (0 < x < 1/(n+1)).$$

Darling gives also the asymptotic joint distribution

$$\lim_{n \to \infty} P\left\{\frac{x}{(n+1)^2} \leq V_{\min}, \ V_{\max} \leq \frac{y + \log(n+1)}{n+1}\right\} = \exp(-x - e^{-y}) \quad (x > 0, \ -\infty < y < \infty). \tag{4}$$

We learn from this equation that $Vmax$ is asymptotically Gumbel distributed and $Vmin$ is asymptotically exponential. In the notation of Gnedenko (1943), the asymptotic distribution functions of V_{max} and V_{min} are, respectively, $\Lambda(x)$ and $1 - \Psi_\alpha(x)$ with $\alpha = 1$. These results can be explained by the well known fact, that if $E_1, E_2, \dots, E_{n+1}$ are independent unit-exponential and T_{n+1} is their sum, then

$$(V_1, \ V_2, \ \dots, \ V_{n+1}) \overset{D}{=} \frac{(E_1, \ E_2, \ \dots, \ E_{n+1})}{T_{n+1}} = \frac{n+1}{T_{n+1}} \cdot \frac{(E_1, \ E_2, \ \dots, \ E_{n+1})}{n+1},$$

independent of T_{n+1}. Since $T_{n+1}/(n+1) \to 1$ a.s., the limit in (4) follows.

Weiss (1959, 1960) writes that the statistics $R_n = V_{max} - V_{min}$ and $S_n = V_{max} = V_{min}$ have been proposed to test H_0 when the alternative is that the $\{X_i\}$ are iid from some df F ($\neq F_0$). He then shows that (asymptotically) the test based on S_n is equivalent to the test based on V_{min} alone, and the test based on R_n is equivalent to the one based on V_{max} alone (small values of V_{min} are critical, large values of V_{max} are critical). He farther shows that the test based on V_{min} is not consistent while the one based on V_{max} is admissible and consistent under any fixed alternative. This is a good reason to check how well V_{max} performs at the present set-up in comparison to other possible tests.

In all the tests that follow, the significance level is α = .05. In particular, the LST rejects H_0 when $V_{max} > v_n(.95)$, where $v_n(p)$ is the p-quantile of V_{max}, determined by (3). Since our study includes small values of n, we do not rely on the asymptotic distribution.

Kolmogorov-Smirnov distance. One of the most popular tests of the null hypothesis is the Kolmogorov-Smirnov test (KST). Given a sample, we define the empirical cumulative distribution function F_n by

$$F_n(x) = \frac{1}{n} \sum_1^n I\{X_i \leq x\} \quad (0 \leq x \leq 1),$$

where $I\{A\}$ is the indicator of the event A. The Kolmogorov-Smirnov distance is defined by

$$D_n = \sup_{0 \leq x \leq 1} |F_n(x) - x|.$$

The KST rejects the null hypothesis with significance level α = .05 if $D_n > d_n(.95)$, where $d_n(p)$ is the p-quantile of D_n. The exact distribution function of D_n is too complicated to write in a closed form. The asymptotic distribution is given by

$$\lim_{n \to \infty} P\{\sqrt{n}D_n \geq x\} = 2\sum_{i=1}^{\infty} (-1)^{i-1} \exp(-2i^2x^2).$$

The quantiles $d_n(.95)$ are well tabulated for $n \leq 50$. For larger n, it is suggested to use the approximation $d_n(.95) \approx 1.36/\sqrt{n}$. However, we prefer to use estimators of $d_n(.95)$ based on Monte Carlo simulations of 10^6 replications of samples of size n. They give more accurate
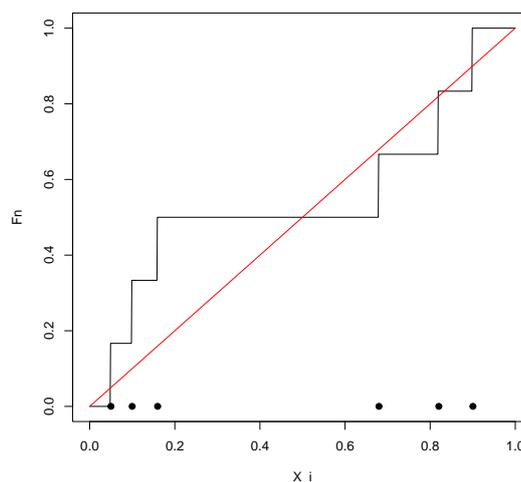
Figure 2. Example of F_n, $n = 6$

results, in the sense that the empirical powers under H_0 are closer to the nominal level of 5%.

Sample serial correlation. As mentioned above, the SSC is the least squares estimator of ρ. Under normality it is also the maximum likelihood estimator (MLE) of ρ (the MLE in our set-up,

with uniform errors, is given in Section 4). The exact distribution is not known but by White (1961), SSC is asymptotically normally distributed with mean $(1 - 2n^{-1} + 4n^{-2})\rho$ and variance equal to $n^{-1}(1 - \rho^2) + n^{-2}(10\rho^2 - 1)$, (ignoring terms of order n^{-3}). Based on SSC, H_0 is rejected when SSC$>$ $s_n(.95)$, the .95-quantile. As in the case of KS, the quantiles were estimated by $N = 10^6$ replications of samples of size n. It turns out that the asymptotic approximations are quite good (relative error $\leq$ 1%) when $n \geq 5000$.

3   POWER COMPARISONS

In this section we present the Monte Carlo results (using the software R) concerning the power of each test. For each selected n and ρ, we generated 10^5 samples of size n according to the autoregressive model (1). The empirical power reported is the proportion of samples for which H_0
H_0 is rejected. At an early stage of our study we thought that LS and KS distance might be highly correlated, because if LS is large, so must be at least one of the adjacent vertical distances, either at the right or at the left of the largest spacing as seen in Figure 2.

Table 1 shows the results for $n = 100$ and 1000. We also included the relative frequency of simultaneous rejection by LST and KST (under *Both*) and their correlation (under $r(V,D)$). It appears that they are indeed quite positively dependent. For $n = 100$ the LST is superior to the KST, inferior to the SSC test. For $n = 1000$, LST is by far superior to the other two tests. To strengthen the impression that as n increases, the superiority of LST becomes more and more apparent, we ran similar simulations for several sample sizes (again 10^5 replications for each pair (n, ρ)). The results are given in tables which appear in the Appendix. Based on these tables, we produced Figure 3. The blue graphs correspond to the likelihood ratio test, which is obviously most powerful (discussed in detail in Section 4). The other three are *ad hoc* tests applied to a particular model. The trend is clear – for $n = 10, 20$ the order (in terms of power) is SSC, KST, LST. For $n = 50, 100$ the order is SSC, LST, KST. For $n \geq 200$ the order is LST, SSC, KST. Moreover, the LST power function, for large n, tends to be very steep and approaches 1 very fast.

Table 1. Empirical Powers, $\alpha = .05$, $n = 100$, 1000

		$n = 100$						$n = 1000$			
ρ	SSC	V_{max}	D_n	Both	$r(V,D)$	ρ	SSC	V_{max}	D_n	Both	$r(V,D)$
0	.050	.050	.050	.006	.221	0	.050	.050	.050	.003	.055
.05	.123	.044	.059	.006	.246	.01	.091	.061	.052	.003	.058
.10	.250	.145	.092	.020	.308	.02	.153	.233	.061	.013	.066
.15	.436	.302	.160	.072	.356	.03	.244	.471	.078	.035	.082
.20	.625	.592	.354	.221	.412	.04	.347	.676	.111	.071	.104
.25	.802	.749	.612	.500	.424	.05	.474	.816	.171	.133	.129
.30	.912	.914	.878	.777	.446	.06	.596	.904	.288	.248	.147
.35	.962	.968	.972	.928	.464	.07	.712	.955	.497	.455	.153
.40	.989	.985	.994	.982	.484	.08	.807	.978	.768	.728	.167
						.09	.884	.993	.940	.921	.187
						.10	.934	.997	.992	.986	.180
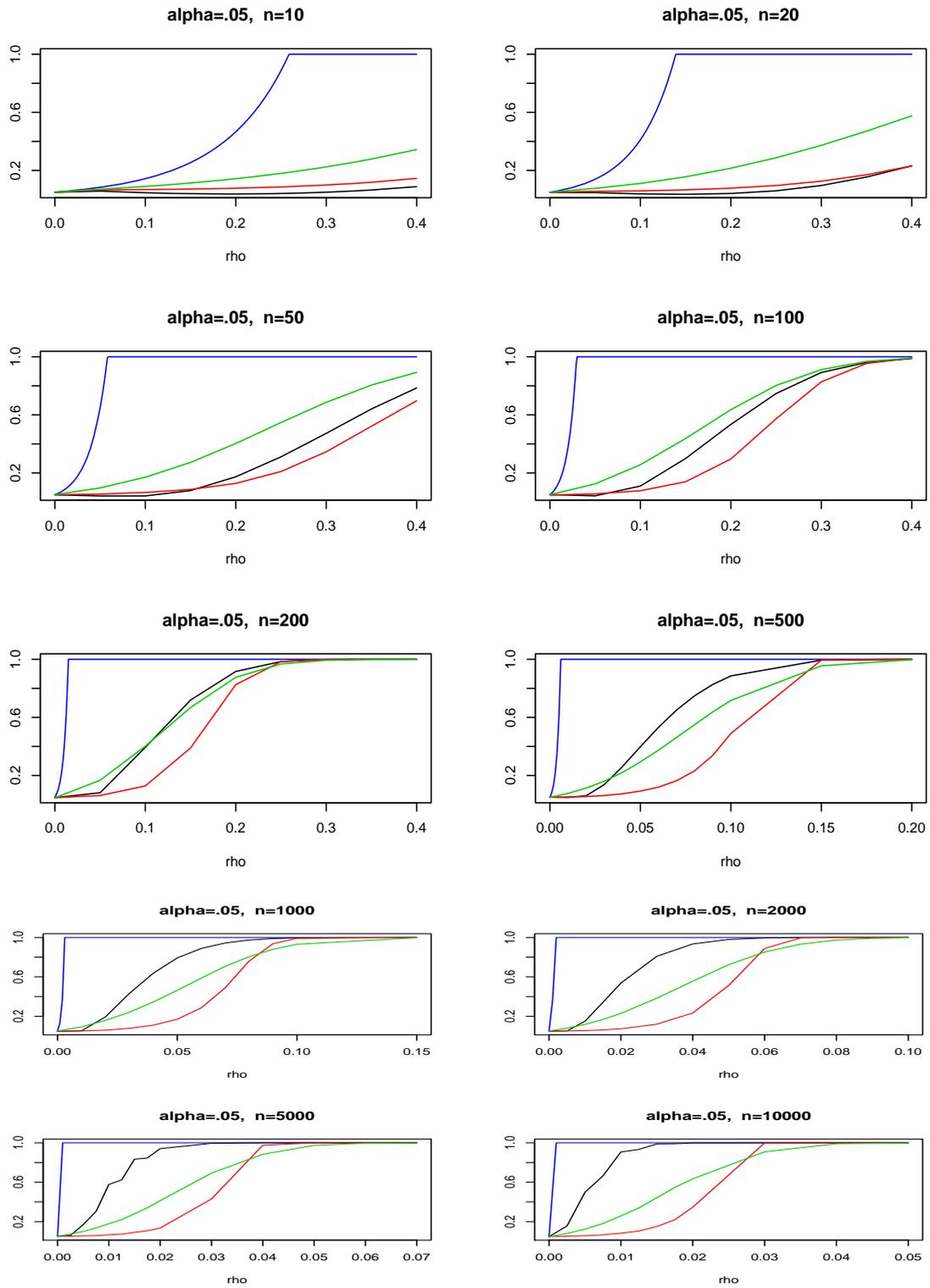
Figure 3. Power Functions, Autoregressive Model, LRT (blue),
LST (black), KST (red), SSC (green)

We can confidently infer that the LST is superior to the KST in the presence of serial correlation. However, under a departure from uniformity, preserving the independence, the KST might be superior to the LST. For example, we took samples of iid $\beta(\gamma, 1)$, namely $X_i = U_i^{1/\gamma}$. Again, for each pair (n, γ), we generated 10^5 samples of size n with the proper γ. The empirical powers are shown in Figure 4. The superiority of the KST over the LST is evident. The KST is quite close to the most powerful test (the blue graph). The latter is the LRT for this model, namely, reject "$\gamma = 1$" when $-2\ \Sigma\ \log X_i < \chi^2_{2n}(.05)$ if the alternative is $\gamma > 1$, or when $-2\ \Sigma\ \log X_i > \chi^2_{2n}(.95)$ if the alternative is "$\gamma < 1$". The power of the LRT is computed directly from the χ^2-distribution, no need for simulations.
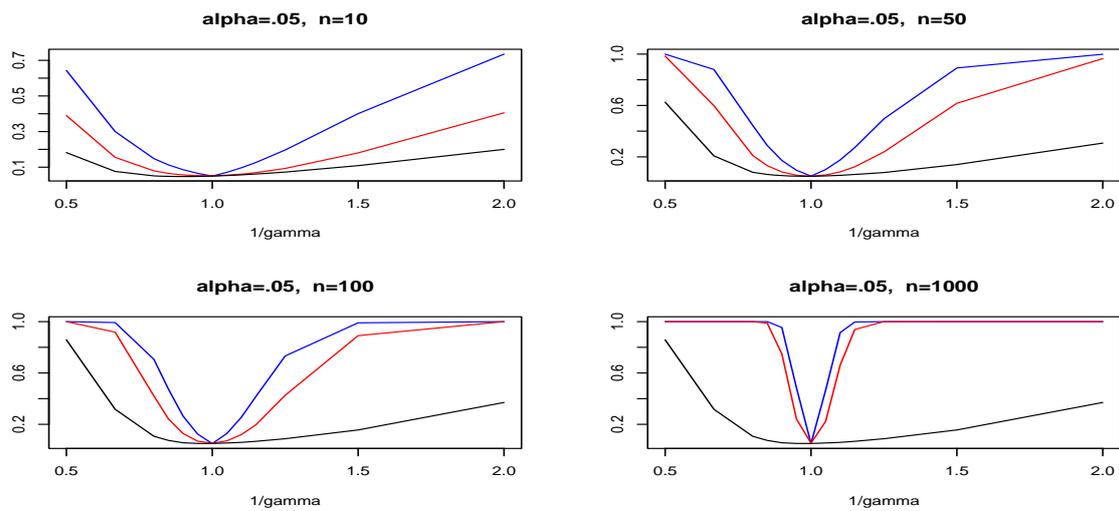
Figure 4. Power Functions, Beta Model, LRT (blue), LST (black), KST (red)

4 LIKELIHOOD-RATIO TEST

In the previous section we presented some evidence in favor of the largest spacing when the alternative is the autoregressive model. However, what is the most powerful test for this alternative? By the Neyman-Pearson theory, the answer is the likelihood-ratio test (LRT).

Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$ be defined by Equation (1). In order to compute the joint density function of $\mathbf{X}$, conditioned on $U_0 = X_0 = x_0$, we express $\mathbf{U} = (U_1, U_2, \dots, U_n)$ in terms of $\mathbf{X}$, namely

$$U_i = (X_i - \rho X_{i-1})\ (1 - \rho)^{-1}\ (1 \le i \le n).$$

The Jacobian of this transformation is $(1 - \rho)^{-n}$. Since $0 \le U_i \le 1$ for all i, one has for $\mathbf{x}$ in $[0, 1]^n$

$$f_X(x) = (1-\rho)^{-n} \, \Pi_{i=1}^n \quad I\{\rho x_{i-1} \le x_i \le \rho x_{i-1} + 1 - \rho\}$$

$$= (1-\rho)^{-n} I\left\{\rho \le \min_{1\le i \le n} \quad \min\left\{\frac{x_i}{x_{i-1}}, \ \frac{1-x_i}{1-x_{i-1}}\right\}\right\}. \tag{5}$$

Let

$$T_{1i} = \min\left\{\frac{X_i}{X_{i-1}}, \ \frac{1-X_i}{1-X_{i-1}}\right\} \quad \left(1 \le i \le n, \ T_{1\min} = \min_{1\le i \le n} T_{1i}\right),$$

then the following facts follow from Equation (5):

Fact 1. *The $\{T_{1i}\}$ are iid uniform on [ρ, 1].*

Fact 2. *The likelihood function is given by*

$$L(\rho) = (1 - \rho)^{-n} I\{\rho \le T_{1min}\} \quad (0 \le \rho \le 1). \tag{6}$$

Fact 3. *The statistic T_{1min} is sufficient with respect to ρ and it is the maximum likelihood estimator (MLE) of ρ.*

For testing
$$H_0 : \rho = 0 \text{ vs. } H_1 : \rho > 0 ,$$

the likelihood function is also the likelihood-ratio. A most powerful α-level test rejects *H*0 when

$$T_{1min} > c_\alpha = 1 - \alpha^{1/n} \tag{7}$$

and the power is given by

$$\pi_\alpha(\rho) = \begin{cases} \dfrac{\alpha}{(1-\rho)^n} & \text{if } \rho \le c_\alpha, \\[4mm] 1 & \text{if } \rho \le c_\alpha. \end{cases}$$

So, in fact, the LRT is based on the minimum of 2*n* ratios. The blue graphs in Figure 3 represent the power of the LRT. No doubt the LRT is superior to all other tests, but we note its close proximity to the LST graph for large *n*.

If one desires $\pi_\alpha(\rho) = 1$ for $\rho \ge \rho_0$, one needs

$$n \ge \frac{\log \alpha}{\log(1 - \rho_0)}. \tag{8}$$

Table 2 presents the (smallest) required sample size for $\alpha = .05$.

Table 2. Sample Size Required for Power 1, $\alpha = .05$

ρ_0	$n \geq$
.1	29
.01	299
.001	2995
10^{-m}	3×10^m

Since $nc_\alpha = -\log \alpha + O(n^{-1})$, for large n, $\pi_\alpha(\rho) = 1$ for $\rho \geq (-\log \alpha)=n$.

5 AUTOREGRESSIVE MODEL OF ORDER k

Suppose we want to protect ourselves against serial correlation of higher order. For this purpose we assume that the data have been generated by an autoregressive model of order k, namely

$$X_i = \rho_1 X_{i-1} + \rho_2 X_{i-2} + \ldots + \rho_k X_{i-k} + \eta U_i \ (1 \leq i \leq n), \tag{9}$$

where $\{U_i : i \geq -k+1\}$ is a $U(0, 1)$-iid sequence and

$$Xi = U_i \ \text{ for } i \leq 0, \ \rho j \geq 0, \ 0 < \eta = 1 - \sum_{j=1}^{k} p_j \leq 1.$$

The goal is to test whether $\rho j = 0$ for all j vs. the alternative that $\rho j > 0$ for at least one j, namely,

$$H_0 : \max_j \rho_j = 0 \ \text{ vs. } \ H_1 : \max_j \rho_j > 0 .$$

The joint density function of $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ at $\mathbf{x} \in [0, 1]^n$, conditioned on $\{X_i = x_i : -k+1 \leq i \leq 0\}$ is given by

$$f_X(x) = \eta^{-n} \prod_{i=1}^{n} I\left\{ \Sigma_{j=1}^{k} \rho_j x_{i-j} \leq x_i \leq \Sigma_{j=1}^{k} \rho_j x_{i-j} + \eta \right\} . \tag{10}$$

Unfortunately, a sufficient statistic of low dimension, as in the case $k = 1$, does not exist. However, if we define

$$T_{ji} = \min\left\{ \frac{X_i}{X_{i-j}}, \frac{1-X_i}{1-X_{i-j}} \right\} \ (1 \leq i \leq n), \ T_{j\min} = \min_{1 \leq i \leq n} T_{ji} ,$$

Then, with probability 1, $\rho_j \leq T_{j\min}$ for $1 \leq j \leq k$.

Hence, it is reasonable to use $T^* = \max_{1 \leq j \leq k} T_{jmin}$ as a test statistic, i.e., to reject $H0$ when $T^* > C_\alpha = t_{kn}^*(1-\alpha)$ - quantile of T^* under H_0. We note that under H_0, within each vector $\boldsymbol{T}_j = (T_{j1}, \ldots, T_{jn})$, the T_{ji} are iid $U(0, 1)$ random variables. Moreover, T_{ji} and $T_{j'i'}$ are independent, except in the case where $i = i'$, implying that the T_{jmin} are dependent. This is why the derivation of c_α in a closed form is too complicated and we must resort to Monte Carlo methods. However, once we

determine c_α for given k and n, we can guarantee power 1 for all alternatives with max $\rho_j \geq c_\alpha$. In order to compute $c_{.05}$ we generated 10^6 samples for each combination of $n = 100, 1000, 10000$ and $k = 1, 2, \ldots, 10$. Here we report, in Table 3, the asymptotic values of $nc_{.05}$ with 3 significant digits.

Table 3. Asymptotic Values of $nc_{.05}$

k	$nc_{.05}$	k	$nc_{.05}$
1	3.00	6	4.40
2	3.54	7	4.51
3	3.85	8	4.62
4	4.08	9	4.72
5	4.24	10	4.79

For instance, suppose we plan to run a Monte-Carlo method based on random numbers generated by a specific software. Suppose further that we want to be sure that there is no positive serial correlation up to lag $k = 8$ and that we can tolerate correlations below $\rho_0 = 10^{-4}$. Then, we have to generate a sample of size $n \geq 4.62/\rho_0 = 46200$ in order that $\max_{1 \leq j \leq 8} \rho_j > 10^{-4}$ will be detected with probability 1.

Remark 2. The resources required to carry out this kind of a test are (almost) free. For instance, in S-PLUS or R, one command does the job: "x=runif(46200)" and it takes a fraction of a second.

Remark 3. A special case of our model is when it is known that $\rho_j = 0$ for $1 \leq j \leq k-1$ for some k and we want to test whether $\rho_k = 0$. Then a test based on T_{kmin} has all the properties as the test based on T_{1min} discussed in Section 3.

Although it is hard to compete with T^*, we still wish to explore the performance of the tests considered in Section 2. Here, instead of the sample serial correlation SSC, we use the sum of squares for regression (SSREG) as the test statistic. Namely, $\mathbf{X} = (X_1, \ldots, X_n)$ is the dependent variable and $\mathbf{X}_j = (X_{1-j}, \ldots, X_{n-j})$, $1 \leq j \leq k$, are the regressors. In Table 4 we present some Monte Carlo results for $k = 2$, based on 10^5 replications of samples of size $n = 100$.

The conclusion is similar to the one of the case $k = 1$. Beside the clear dominance of T^*, we note that the performance of SSREG is very poor. The reason could be that the deviates from the model are uniform rather than normal. The largest spacing test performs better than the other two.

Table 4. Empirical Powers, $k = 2$, $\alpha = .05$, $n = 100$

ρ_1	ρ_2	$SSREG$	V_{max}	D_n	T^*
0	0	.050	.050	.050	.050
.025	0	.049	.042	.055	.468
0	.025	.049	.042	.056	.459
.025	.025	.047	.045	.063	.990
.050	.050	.056	.171	.090	1.000
.075	0	.061	.072	.070	1.000
0	.075	.070	.072	.070	1.000
.075	.075	.076	.565	.171	1.000
.100	0	.079	.146	.090	1.000
0	.100	.090	.150	.091	1.000
.100	.100	.107	.848	.379	1.000
.125	0	.111	.245	.121	1.000
0	.125	.113	.250	.121	1.000
.125	.125	.146	.960	.738	1.000
.150	.150	.197	.993	.953	1.000

6 TWO MORE MODELS

We wish we could claim that the LST is a powerful test to detect dependence in general. For this, one has to examine its performance against all kinds of dependencies – an impossible mission. We have experimented with several models with serial dependence and the conclusion is similar – except for very small sample sizes, LST is more powerful than KST. Here are two examples.

Binomial model. Let U_0, U_1, U_2, ... be iid uniform as before and let B_1, B_2, ... be a Bernoulli sequence with parameter p, independent of the U-sequence. The *binomial sequence Y* is defined by $Y_i = B_i Y_{i-1} + (1 - B_i)U_i$ ($i \geq 1$, $Y_0 = U_0$). The marginal distribution of each Y_i is U [0, 1], the first serial correlation is p and so is $P\{Y_i = Y_{i+1}\} = p$. Clusters of equal neighbors are of random (geometric) length (see Figure 5 for a scatter points (i, Y_i)).

Moving-max model. Let ξ_1, ξ_2, ... be a sequence of iid $\beta(k^{-1}, 1)$ random variables, where k is a fixed positive integer. Let $Z_i = \max\{\xi_i, \xi_{i+1}, ..., \xi_{i+k-1}\}$ ($i \geq 1$). The Z-sequence is called *a moving-max sequence of order k*. For each i, Z_i is U [0, 1]-distributed but neighboring values are dependent. Upper extreme values appear in clusters of size k, which imply that the *extremal index* is equal to k^{-1}. For $k = 2$, the first serial correlation is 3/7 and $P\{Z_i = Z_{i+1}\} = 1/3$.

The two plots in Figure 5 look very similar. In both cases, the experienced practitioner will reject the independence hypothesis just on the basis of the fact that for continuous random variables, the probability of a tie is 0. We brought these cases to see how well the LST and the KST detect the dependence.
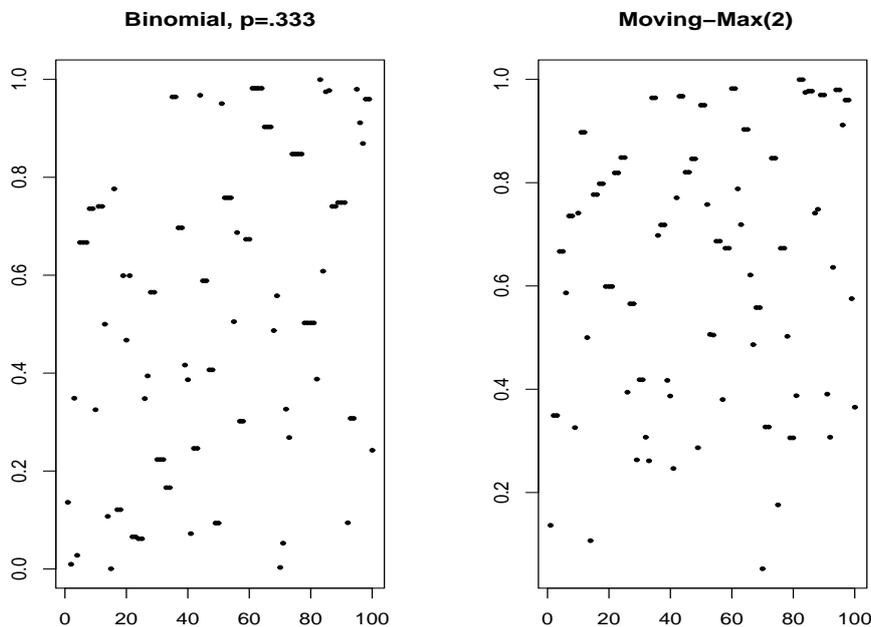
Figure 5. Scatter plot for the two models

Figure 6 shows the (empirical) power functions of the two tests applied to the moving-max models of order $k = 2$ and $k = 3$. The superiority of LST over KST is evident, moreover, LST is even not consistent in this case. A similar Monte Carlo study was carried out on the binomial model. The results are shown in the Appendix. The conclusion is very similar to the conclusion

regarding the autoregressive model, i.e., for $n \geq 100$, LST is superior to the KST, and as n increases, it becomes more and more so.
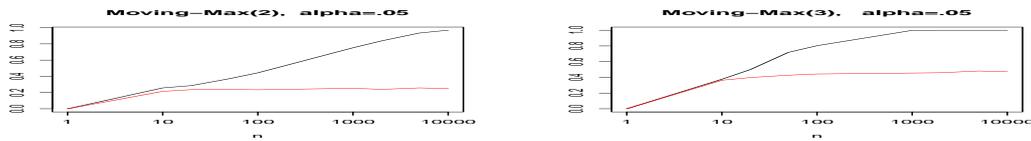
Figure 6. Power Functions, Moving-Max Model (logarithmic scale), LST (black), KST (red)

Remark 4. The binomial model has some resemblance to the autoregressive model (where ρ is replaced by a random variable with mean p). Suppose we apply the LRT of the autoregressive model to a sample from the binomial model with some $p > 0$. Then, for all i such that $B_i = 0$, the T_{1i} are (as in Section 4) iid uniform on $[0, 1]$. For all i such that $B_i = 1$, one has $T_{1i} = 1$. Hence, given $S = \sum_1^n (1 - B_i)$,

$$P\{T_{1min} > 1 - \alpha^{1/n}\} = \alpha^{S/n}.$$

Since $S/n \to q = 1 - p$ a.s. as $n \to \infty$, the power tends to $\alpha^q (= .0675$ for $\alpha = .05$ and $p = .1)$. Applying the same test to data from the moving-max model yields even lower power.

7 CONCLUSIONS

The main theme of the paper is to show that the largest spacing of a sample is sensitive to serial correlation and is quite powerful in detecting it, more powerful than the Kolmogorov-Smirnov distance. The opposite is true under independence but the true distribution is diœerent from the null distribution. Of course, when the data are generated by a specific (known) parametric model, and there exists a most powerful test, one should use that test. We went into detail in Sections 4 and 5 because we found it interesting to learn (for the first time in the statistical literature, to the best of our knowledge) that the MLE of a serial correlation is a (lower) sample extreme, and a test based on it is most powerful. The overall message is clear. if you worry about serial dependence in your data and you cannot assume a particular model, the LST is a reasonable test to use.

8 REFERENCES

Darling, D.A. (1953). On a class of problems related to the random division of an interval. *Ann. Math. Statist.* **24**: 239-253.

Gentle, J.E. (2003). *Random Number Generation and Monte Carlo Methods*. New York: Springer.

Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**: 423-453.

Lawrance, A.J. (1992). Uniformly distributed first-order autoregressive time series models and multiplicative congruential random number generators. *J. Appl. Prob.* **29**: 896-903.

L'Ecuyer, P. & Simard, R. (2007). TestU01. A C library for empirical testing of random number generators. *ACM Transactions on Mathematical Software.* **33**, No. 4, Article 22.

Onn, S. & Weissman, I. (2011). Generating uniform random vectors over a simplex with implications to the volume of a certain polytope and to multivariate extremes. *Annals of Operations Research.* **189**: 331-342.

Weiss, L. (1959). The limiting joint distribution of the largest and smallest sample spacings. *Ann. Math. Statist.* **30**: 590-593.

Weiss, L. (1960). A test of fit based on the largest sample spacing. *J. Soc. Indust. Appl. Math.* **8:** 295-299.

White, J.S. (1961). Asymptotic expansions for the mean and variance of the serial correlation coefficient. *Biometrika* **48**: 85-94.

Whitworth, W.A. (1897). *Choice and Chance.* Cambridge Univesity Press.

APPENDIX

The empirical powers of the Monte Carlo simulations for the autoregressive model are given below. Each entry is a result of 10^5 replications. The graphs in Figure 3 are based on Table 5.

Table 5. Empirical Powers, Autoregressive Model

	$n = 10$				$n = 20$		
ρ	SSC	V_{max}	D_n	ρ	SSC	V_{max}	D_n
0.00	0.05133	0.04975	0.05037	0.00	0.05051	0.04971	0.05038
0.05	0.07103	0.05769	0.06590	0.05	0.07656	0.04728	0.05602
0.10	0.08995	0.04648	0.06801	0.10	0.11030	0.03834	0.05934
0.15	0.11316	0.03983	0.07153	0.15	0.15607	0.03559	0.06609
0.20	0.14359	0.03763	0.07762	0.20	0.21530	0.04134	0.07790
0.25	0.17958	0.04137	0.08681	0.25	0.28692	0.05967	0.09602
0.30	0.22447	0.0499	0.09939	0.30	0.37274	0.09604	0.12600
0.35	0.27866	0.06514	0.11846	0.35	0.47051	0.15446	0.17001
0.40	0.34355	0.08843	0.14538	0.40	0.57579	0.23172	0.23157

	$n = 50$				$n = 100$		
ρ	SSC	V_{max}	D_n	ρ	SSC	V_{max}	D_n
0.00	0.05043	0.04969	0.05023	0.00	0.04979	0.05106	0.04822
0.05	0.09709	0.04166	0.05470	0.05	0.12491	0.04251	0.05587
0.10	0.17079	0.04192	0.06556	0.10	0.25663	0.10982	0.07755
0.15	0.27259	0.07875	0.08652	0.15	0.43726	0.29966	0.13949
0.20	0.40277	0.17347	0.12812	0.20	0.63583	0.53433	0.29601
0.25	0.54738	0.31108	0.21007	0.25	0.80235	0.74673	0.57287
0.30	0.68625	0.47260	0.34568	0.30	0.91155	0.89186	0.82717
0.35	0.80513	0.64069	0.52163	0.35	0.96809	0.96080	0.95418
0.40	0.89275	0.78542	0.69733	0.40	0.99113	0.98766	0.99162

	n = 200				n = 500		
ρ	SSC	V_{max}	D_n	ρ	SSC	V_{max}	D_n
0.00	0.04640	0.04720	0.04740	0.00	0.04912	0.05077	0.05007
0.05	0.16710	0.08130	0.06220	0.01	0.07674	0.04870	0.05142
0.10	0.40150	0.39340	0.12790	0.02	0.11337	0.06056	0.05484
0.15	0.66820	0.71950	0.38950	0.03	0.16028	0.13689	0.06191
0.20	0.87580	0.91640	0.82570	0.04	0.22103	0.26020	0.07383
0.25	0.96790	0.98460	0.98410	0.05	0.29361	0.39809	0.09249
0.30	0.99340	0.99730	0.99890	0.06	0.37480	0.53170	0.11984
0.35	0.99930	0.99980	1.00000	0.07	0.46123	0.65072	0.16376
0.40	0.99990	0.99990	1.00000	0.08	0.54949	0.74919	0.23214
				0.09	0.63764	0.82675	0.33809
				0.10	0.71659	0.88551	0.48909
				0.15	0.95492	0.99416	0.99432
				0.20	0.99754	0.99950	1.00000

	n = 1000				n = 2000		
ρ	SSC	V_{max}	D_n	ρ	SSC	V_{max}	D_n
0.00	0.05050	0.04840	0.04750	0.000	0.05111	0.04949	0.04976
0.01	0.09250	0.05280	0.05080	0.005	0.07938	0.05189	0.05127
0.02	0.15840	0.19390	0.05910	0.010	0.11750	0.15014	0.05435
0.03	0.24100	0.43230	0.07820	0.015	0.16819	0.34179	0.06163
0.04	0.34580	0.63770	0.10970	0.020	0.22958	0.53729	0.07315
0.05	0.46370	0.79240	0.17040	0.030	0.38317	0.80883	0.11937
0.06	0.58670	0.88840	0.28550	0.040	0.55600	0.93420	0.23348
0.07	0.70690	0.94400	0.48920	0.050	0.72345	0.98030	0.51278
0.08	0.80440	0.97450	0.76100	.060	0.85195	0.99537	0.88792
0.09	0.88040	0.98900	0.93690	0.070	0.93184	0.99915	0.99573
0.10	0.93130	0.99680	0.99170	0.080	0.97397	0.99976	0.99998
0.15	0.99780	0.99990	1.00000	0.090	0.99178	0.99998	1.00000
				0.100	0.99759	1.00000	1.00000

	n = 5000				n = 10000		
ρ	SSC	V_{max}	D_n	ρ	SSC	V_{max}	D_n
0.0000	0.0509	0.0489	0.0548	0.0000	0.0471	0.0468	0.0481
0.0025	0.0741	0.0589	0.0515	0.0025	0.0800	0.1566	0.0519
0.0050	0.0994	0.1694	0.0554	0.0050	0.1235	0.4980	0.0567
0.0075	0.1356	0.3035	0.0578	0.0075	0.1810	0.6642	0.0654
0.0100	0.1782	0.5778	0.0661	0.0100	0.2580	0.9067	0.0827
0.0125	0.2204	0.6239	0.0724	0.0125	0.3382	0.9331	0.1044
0.0150	0.2794	0.8338	0.0907	0.0150	0.4434	0.9888	0.1505
0.0175	0.3406	0.8456	0.1070	0.0175	0.5470	0.9908	0.2182
0.0200	0.4131	0.9411	0.1350	0.0200	0.6345	0.9988	0.3466
0.0300	0.6922	0.9949	0.4299	0.0300	0.9093	1.0000	0.9982
0.0400	0.8838	0.9999	0.9754	0.0400	0.9908	1.0000	1.0000
0.0500	0.9744	1.0000	1.0000	0.0500	0.9997	1.0000	1.0000
0.0600	0.9961	1.0000	1.0000				
0.0700	0.9997	1.0000	1.0000				

The following plots show the empirical powers of the Binomial Model, based on 10^4 replications for each pair (*n, p*).
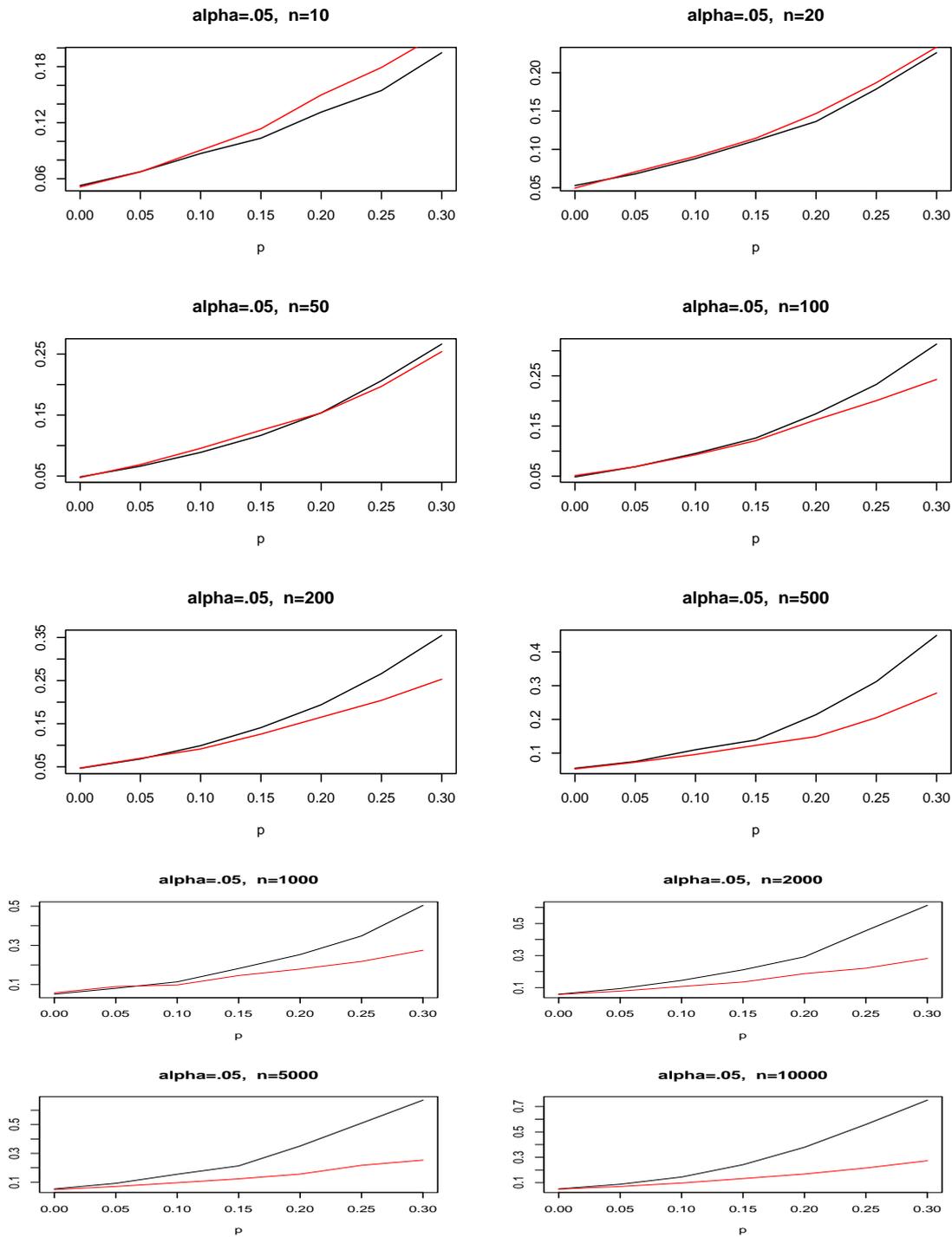
Figure 7. Power Functions, Binomial Model, LST (black), KST (red)