Exact Kolmogorov-Smirnov Test of Normality
for Completely Randomized Designs

by

Michele M. Altavela and Constance L. Wood

Summary

Considered are the finite-sample properties of the Kolmogorov-Smirnov test statistics for testing normality of errors in completely randomized designs. Monte Carlo simulations for the critical values 2, 10(2) treatments and 3, 4, 5, 10, 15, 20 replications per treatment are given. A comparison is made with the corresponding results of Lilliefors (1967) for the one-sample test statistics for normality with estimated parameters.

Paper No. BU-609-M in the Biometric Unit Series, Cornell University, Ithaca, New York 14853.

This research was supported by Hatch Project 402.

Key Words: Goodness-of-fit; Completely randomized design; Monte Carlo simulation; Normality.

1. Introduction.

A frequently occurring problem is that of testing the validity of the generally assumed linear model for data arising from a completely randomized design, i.e.,

$$(1.1) \qquad Y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \qquad j = 1, \cdots, n_i \quad \text{and} \quad i = 1, \cdots, a \ ,$$

where α_i is the non-stochastic effect of the i^{th} treatment with $\sum_{i=1}^{a} n_i \alpha_i = 0$ and ϵ_{ij}, the experimental error associated with the j^{th} experimental unit in the i^{th} treatment group, is normally distributed with zero mean and unknown variance σ^2 .

The relevant departures from these assumptions are (i) a non-additive error structure resulting in heteroscedasticity of the experimental errors and (ii) non-normality of the experimental errors. If both (i) and (ii) are of concern, then Serfling and Wood (1975) suggest basing a test of the hypothesis that (1.1) is the correct model on the combined modified empirical cumulative distribution function (c.d.f.)

$$H_{n.}(t) = (n_.)^{-1} \sum_{i=1}^{a} \sum_{j=1}^{n_i} I\left[\Phi\left(\frac{Y_{ij} - \bar{Y}_{i.}}{s_p}\right) \le t\right], \qquad 0 \le t \le 1 \ ,$$

where $I(A)$ is the indicator of the event A, Φ is the standard normal c.d.f., $n_. = \sum_{i=1}^{a} n_i$ and s_p^2 is the usual pooled variance estimator of σ^2 . However, if only (ii) is relevant, the appropriate modified c.d.f. is

$$H_{n.}(t) = n_.^{-1} \sum_{i=1}^{n} I\left[\Phi\left(\frac{Y_{ij} - \bar{Y}_{i.}}{s_i}\right) \le t\right] \ ,$$

where s_i^2 is now the sample variance for the i^{th} treatment group. (We will employ the usual dot notation in the remainder of this paper.) In both cases it is shown in Serfling and Wood (1975) that the associated Kolmogorov-Smirnov statistics

(1.2) $$n_{\cdot}^{\frac{1}{2}} D_n^+ = \sup_{0\leq t\leq 1} (n_{\cdot})^{\frac{1}{2}}[H_{n.}(t) - t] \; ,$$

(1.3) $$n_{\cdot}^{\frac{1}{2}} D_n^- = \inf_{0\leq t\leq 1} (n_{\cdot})^{\frac{1}{2}}[H_{n.}(t) - t] \; ,$$

and

(1.4) $$n_{\cdot}^{\frac{1}{2}} D_n = \sup_{0\leq t\leq 1} (n_{\cdot})^{\frac{1}{2}}|H_{n.}(t) - t|$$

have limiting distributions which coincide with those of the one-sample Kolmogorov-Smirnov statistics with estimated mean and variance. Here we consider the finite sample properties of these statistics under the null hypothesis (1.1). Monte Carlo simulations of the critical values are given in Section 2. A comparison with the one-sample results of Lilliefors (1967) is made in Section 3.

2. Results. The exact finite sample size distributions for the Kolmogorov-Smirnov statistic were simulated using both the pooled and the individual treatment variance estimators. Since

(2.1) $$\left(Y_{ij} - \bar{Y}_{i.}\right)/s_i, \qquad i = 1, \cdots, a; \quad j = 1, \cdots, n$$

and

(2.2) $$\left(Y_{ij} - \bar{Y}_{i.}\right)/s_p, \qquad i = 1, \cdots, a; \quad j = 1, \cdots, n_i$$

are location and scale invariant, only standard normal variables were generated. Then (2.1), (2.2), and the respective Kolmogorov-Smirnov statistics were generated. One thousand simulations were run for a = 2, 10(2) and $n_i = n = 3, 4, 5, 10, 15, 20$. The standard normal variables were generated using the double precision version of the International Mathematical and Statistical Library (IMSL) subroutine GGNOR. The results for the pooled variance estimator of σ^2 are given in Table 1 and Table 2.

TABLE 1.

Simulated Quantiles for D_n with Estimated Parameters $\left(\bar{X}, s_p^2\right)$

		Number of Replications per Treatment					
Quantile		3	4	5	10	15	20
q	a						
.80	2	.25820	.23335	.21290	.15766	.13265	.11422
	4	.19452	.17187	.15833	.11273	.09496	.08216
	6	.16451	.14645	.13310	.09436	.07718	
	8	.14596	.13037	.11653	.08344		
	10	.13803	.11800	.10624	.07497		
	12	.12825	.10990	.09882			
.85	2	.27240	.24345	.22290	.16408	.13768	.12143
	4	.20396	.18048	.16537	.11883	.10023	.08626
	6	.17331	.15434	.13775	.09857	.08140	
	8	.15323	.13595	.12297	.08767		
	10	.14313	.12403	.11076	.07785		
	12	.13269	.11369	.10287			
.90	2	.28730	.25735	.23784	.17379	.14643	.12858
	4	.21421	.19151	.17180	.12731	.10689	.09084
	6	.18167	.16374	.14560	.10504	.08719	
	8	.16284	.14208	.13096	.09356		
	10	.15035	.13220	.11670	.08259		
	12	.13950	.12042	.10935			
.95	2	.30479	.27823	.25809	.18854	.15984	.14117
	4	.22785	.20810	.18807	.13735	.11727	.09977
	6	.19668	.17430	.15757	.11322	.09499	
	8	.17375	.15218	.14385	.10293		
	10	.16026	.14443	.12551	.09011		
	12	.15042	.13096	.11846			
.975	2	.31809	.29295	.27530	.20861	.17186	.14963
	4	.24095	.21911	.20446	.14499	.12772	.10687
	6	.20757	.18622	.16888	.12245	.10170	
	8	.18273	.16468	.15256	.10949		
	10	.16737	.15400	.13401	.09609		
	12	.16196	.13799	.12667			
.99	2	.33342	.31427	.29705	.22601	.18214	.16345
	4	.26332	.22816	.21927	.15886	.13717	.11861
	6	.22066	.19838	.18475	.12742	.10702	
	8	.19756	.17562	.16550	.11731		
	10	.18655	.16560	.14172	.10464		
	12	.17090	.14782	.13333			
.995	2	.34556	.32448	.30844	.23770	.20133	.16905
	4	.27742	.24849	.22746	.16946	.13964	.12127
	6	.22637	.20732	.19542	.13667	.11173	
	8	.21079	.17880	.17144	.11987		
	10	.19511	.16887	.14509	.10883		
	12	.18166	.15076	.14053			

TABLE 2.

Simulated Quantiles for D_n^+ with Estimated Parameters $\left(\bar{X}, s_p^2\right)$

		Number of Observations per Treatment					
Quantile		3	4	5	10	15	20
q	a						
.80	2	.23075	.20641	.18938	.14161	.11785	.10073
	4	.17842	.15633	.14156	.10312	.08507	.07220
	6	.15237	.13433	.11790	.08617	.07004	
	8	.13794	.11780	.10449	.07544		
	10	.12843	.10870	.09438	.06841		
	12	.12002	.10218	.08800			
.85	2	.24718	.21870	.20144	.14908	.12454	.10640
	4	.18591	.16385	.14863	.10813	.08977	.07823
	6	.15706	.14233	.12500	.08991	.07357	
	8	.14357	.12392	.11044	.07980		
	10	.13510	.11440	.09992	.07224		
	12	.12464	.10581	.09195			
.90	2	.25993	.23331	.21550	.15707	.13217	.11422
	4	.19876	.17519	.15834	.11677	.09559	.08415
	6	.16932	.15069	.13394	.09428	.07850	
	8	.15238	.13249	.11671	.08411		
	10	.14174	.12088	.10561	.07657		
	12	.13008	.11129	.09872			
.95	2	.28984	.25713	.23955	.17416	.14586	.12769
	4	.21400	.19238	.16904	.12840	.10629	.09215
	6	.18379	.16529	.14642	.10319	.08719	
	8	.16606	.14275	.13065	.09342		
	10	.15377	.13414	.11584	.08424		
	12	.14053	.11912	.10935			
.975	2	.30252	.27860	.25817	.18843	.15675	.14157
	4	.23234	.21131	.18625	.13688	.11727	.10076
	6	.19720	.17522	.15757	.11322	.09522	
	8	.17574	.15612	.14294	.10300		
	10	.16362	.14508	.12214	.09050		
	12	.15053	.12790	.11811			
.99	2	.31894	.29477	.28850	.21018	.17467	.15001
	4	.25425	.21948	.20210	.14691	.13161	.10737
	6	.21302	.18937	.16888	.12392	.10212	
	8	.18664	.16636	.15256	.11377		
	10	.17292	.15784	.13413	.10016		
	12	.16731	.13571	.12667			
.995	2	.33312	.30439	.29953	.22601	.19074	.15682
	4	.27100	.22816	.21320	.15556	.13607	.11730
	6	.22066	.19728	.17976	.12855	.10724	
	8	.19662	.17644	.16059	.11979		
	10	.18655	.16357	.13572	.10550		
	12	.17562	.14160	.13333			

The simulated quantiles using the within-treatment estimators of the experimental error are given in the following tables.

TABLE 3.

Simulated Quantiles for D_n with Estimated Parameters $\left(\bar{X},\ s_i^2\right)$

		Number of Observations per Treatment					
Quantile		3	4	5	10	15	20
q	a						
.80	2	.25985	.23489	.21173	.15582	.13113	.11437
	4	.19060	.16945	.15475	.11371	.09633	.08153
	6	.16508	.14359	.13161	.09398	.07783	
	8	.15133	.12458	.11344	.08174		
	10	.13825	.11509	.10215	.07306		
	12	.13216	.10516	.09383			
.85	2	.27923	.24412	.22104	.16302	.13776	.11986
	4	.20168	.18004	.16301	.11917	.10027	.08531
	6	.17356	.15065	.13652	.09812	.08123	
	8	.15886	.13053	.11925	.08540		
	10	.14323	.12115	.10619	.07621		
	12	.13549	.10965	.09807			
.90	2	.30080	.25801	.23614	.17309	.14625	.12897
	4	.21809	.19169	.17086	.12672	.10528	.09057
	6	.18673	.16152	.14555	.10461	.08600	
	8	.16613	.13857	.12703	.09052		
	10	.15125	.12789	.11344	.08074		
	12	.14411	.11789	.10437			
.95	2	.32874	.27812	.26015	.18827	.15811	.13994
	4	.23997	.20973	.18545	.13816	.11711	.09840
	6	.20278	.17359	.15563	.11206	.09712	
	8	.18339	.15074	.13854	.09939		
	10	.16398	.13892	.12204	.08826		
	12	.15752	.12668	.11225			
.975	2	.34140	.29440	.28010	.20418	.17207	.14887
	4	.26521	.22043	.20257	.14378	.12759	.10703
	6	.22470	.18763	.17019	.12153	.10385	
	8	.20319	.16199	.15185	.10574		
	10	.18096	.14769	.13008	.09509		
	12	.16856	.13741	.11962			
.99	2	.35725	.31091	.29518	.23090	.18399	.16199
	4	.28108	.23832	.22292	.16276	.13711	.11506
	6	.25163	.20215	.17876	.13023	.10763	
	8	.21902	.17472	.16090	.11414		
	10	.19148	.15862	.13669	.10208		
	12	.17909	.14677	.13131			
.995	2	.36458	.32904	.30604	.23627	.19572	.16958
	4	.29370	.25251	.23283	.17347	.14175	.11907
	6	.26485	.20833	.19411	.13339	.10993	
	8	.22550	.17926	.17757	.11863		
	10	.20694	.16525	.14193	.11009		
	12	.19261	.14979	.13970			

TABLE 4.

Simulated Quantiles for D_n^+ with Estimated Parameters $\left(\bar{X}, s_i^2\right)$

		Number of Observations per Treatment					
Quantile		3	4	5	10	15	20
q	a						
.80	2	.21760	.20414	.18831	.14222	.11570	.10095
	4	.17780	.15205	.13823	.10077	.08489	.07237
	6	.15464	.13050	.11668	.08481	.06992	
	8	.13673	.11441	.10065	.07489		
	10	.13059	.10523	.09094	.06632		
	12	.12829	.09644	.08454			
.85	2	.23532	.21854	.19931	.14836	.12419	.10682
	4	.18672	.15953	.14494	.10734	.09018	.07711
	6	.16244	.13787	.12334	.08839	.07339	
	8	.14546	.11941	.10622	.07961		
	10	.13354	.11042	.09544	.06988		
	12	.13019	.10026	.08919			
.90	2	.25985	.23410	.21310	.15654	.13196	.11488
	4	.19917	.17281	.15475	.11423	.09750	.08336
	6	.17245	.14793	.13083	.09398	.07824	
	8	.15572	.12781	.11617	.08320		
	10	.14184	.11847	.10200	.07369		
	12	.13358	.10670	.09533			
.95	2	.30295	.25628	.23629	.17222	.14655	.12833
	4	.22467	.19181	.16923	.12703	.10561	.09118
	6	.19113	.15963	.14274	.10333	.08512	
	8	.16936	.13935	.12836	.09177		
	10	.15652	.12923	.11281	.08136		
	12	.14790	.11621	.10477			
.975	2	.32874	.27892	.26257	.18777	.15811	.13994
	4	.24573	.21058	.18124	.13806	.11656	.09792
	6	.20685	.17296	.15444	.11239	.09702	
	8	.18812	.15338	.13882	.10078		
	10	.17145	.13804	.12238	.08788		
	12	.15811	.12612	.11110			
.99	2	.34752	.30078	.28834	.20060	.17536	.15306
	4	.27151	.22506	.19411	.14718	.13171	.10939
	6	.22991	.18768	.17110	.12574	.10509	
	8	.21516	.16493	.15293	.11003		
	10	.18797	.14769	.13138	.09624		
	12	.17311	.14024	.12211			
.995	2	.35751	.30843	.29431	.23090	.19038	.15817
	4	.28108	.23832	.21492	.16003	.13590	.11885
	6	.25634	.19221	.17876	.13023	.10763	
	8	.22090	.17244	.16470	.11557		
	10	.19614	.15691	.13435	.10487		
	12	.18007	.14682	.13151			

The accuracy of these estimated quantiles was investigated by rerunning the simulation for a = 2 and n = 4.5 . The two independent estimates of the quantiles are shown below:

TABLE 5.

Estimates of Quantiles for Two Treatments

	D_n				D_n^+			
	$(\bar{X}, s_p^2)$		$(\bar{X}, s_i^2)$		(X, s_p^2)		(X, s_i^2)	
Quantile	4	5	4	5	4	5	4	5
.8	.23472 .23197	.21636 .20943	.23613 .23364	.21313 .21033	.20455 .20826	.19151 .18724	.20439 .20388	.18863 .18799
.85	.24297 .24392	.22553 .22026	.24667 .24156	.22082 .22125	.21668 .22072	.20388 .19899	.22156 .21551	.20011 .19851
.90	.25539 .25881	.24007 .23561	.26066 .25536	.23701 .23527	.23142 .23520	.22156 .20943	.23549 .23270	.21575 .21044
.95	.27669 .27977	.25882 .25735	.27963 .27661	.26134 .25895	.25131 .26292	.24242 .23668	.25656 .25599	.23785 .23473
.975	.28994 .29595	.27535 .27524	.29536 .29344	.27998 .28022	.27341 .28378	.25946 .25687	.27828 .27955	.26216 .26298
.990	.30889 .31964	.29705 .29705	.31109 .31073	.29763 .29273	.28946 .30007	.28942 .28758	.29825 .30330	.29467 .28200
.995	.32066 .32829	.31423 .30264	.31835 .33973	.30670 .30537	.29982 .30896	.30124 .29781	.30843 .30843	.30300 .28561

Obviously the .80 quantiles are more precise. Except for a few values, these estimates are consistent in the second decimal place.

3. Comparison with One-Sample Statistics. Lilliefors (1967) published a Monte Carlo simulation of the quantiles of the one-sample Kolmogorov-Smirnov test, $B_{n.}$, for normality with the parameters estimated by $\bar{X}$ and s^2 . In this section, we investigate the relationship between those values and the estimated quantiles given in Section 2. For the sake of comparison, $n_.$, the total number of observations irrespective of treatment, will correspond to the sample size of the one-sample statistic. Because of the greater stability, we will restrict attention to the .80 quantiles of $D_{n.}$, calculated with

the pooled variance estimator. The observed differences are shown below.

TABLE 6.

Estimated Quantiles of B_n. - Estimated Quantiles of D_n.

n a	3	4	5	10	15
2	.007	.014	.002	.002	.002
4	.004	.001	.002		
6	.001		-.002		
10	.007				

It is interesting to note that these differences certainly are the order of the degree of accuracy of both sets of estimated quantiles. This implies that the difference in covariance structure of the standardized variables does not greatly effect the test statistic. Further, a good approximation to the quantiles of D_n is given by the quantiles of B_n.

References

Lilliefors, H. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Am. Statist. Assoc.* 62, 399-402.

Serfling, R. J. and Wood, C. L. (1975). On null-hypothesis limiting distributions of Kolmogorov-Smirnov type statistics with estimated location and scale parameters. Florida State University Statistics Report M339.