Statistica & Applicazioni - Vol. IX, n. 1, 2011, pp. 63-77

EXACT CRITICAL VALUES OF KOLMOGOROV-SMIRNOV TEST FOR DISCRETE RANDOM VARIABLES

Silvia Facchinetti*
Paola Maddalena Chiodini**

SUMMARY

The nonparametric Kolmogorov-Smirnov goodness of fit test is employed to test if a random sample comes from a specified continuous distribution function. When this hypothesis is not satisfied the test is no longer applicable accurately. In the last years relatively little attention has been paid to the problems of the application of Kolmogorov-Smirnov test for discrete distributions. In this paper we present a survey of the previous works, we propose a procedure to apply the test to discrete random variables and we define the corresponding statistic. Moreover for some given distributions the exact critical values are tabulated and a comparison with the continuous case is made.

Keywords: *Goodness of Fit Test, Discrete Distributions, Empirical Distribution Function.*

1. INTRODUCTION

The compatibility of a simple random sample of data with a given distribution can be checked by a goodness of fit tests, which test the null hypothesis against a generic alternative:

$$\begin{cases} H_0 : F(x) = F_0(x) & \text{for all } x \\ H_1 : F(x) \neq F_0(x) & \text{for some } x \end{cases} \tag{1}$$

where $F(x) = P(X \leq x)$ is the true cumulative distribution function of X, and $F_0(x)$ is some hypothesized cumulative distribution function with all parameters specified.

One of the most important goodness of fit test has been proposed by Kolmogorov (1933) and Smirnov (1939).

Let X be a continuous random variable with distribution function $F(x)$, and let $(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ be an ordered simple random sample of size n from X, so that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.

* Dipartimento di Scienze Statistiche - Università Cattolica del Sacro Cuore - L.go Gemelli, 1 - 20123 MILANO (e-mail: ✉ silvia.facchinetti@unicatt.it).
** Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali - Università degli Studi di Milano-Bicocca - P.zza dell'Ateneo Nuovo, 1 - 20126 MILANO (e-mail: paola.chiodini@unimib.it).

We define the empirical distribution function by

$$S_n(x) = \begin{cases} 0 & \text{for } x < x_{(1)} \\ \dfrac{k}{n} & \text{for } x_{(k)} \leq x < x_{(k+1)} \text{with } k = 1, 2, \ldots, n-1. \\ 1 & \text{for } x \geq x_{(n)}. \end{cases} \tag{2}$$

This is a step function with jumps occurring at the sample values.

The Kolmogorov-Smirnov statistic is based on a comparison between the hypothesized distribution function $F_0(x)$ and the empirical one $S_n(x)$:

$$D_n = \sup_{-\infty < x < \infty} |S_n(x) - F_0(x)|. \tag{3}$$

The critical region to refuse the null hypothesis in (1) is:

$$R = \left\{ D_n : D_n > \frac{d_\alpha}{\sqrt{n}} \right\}$$

where d_α depends only on α.

Since X is a continuous random variable, D_n depends on the probability integral transformation of the sample values and the probability distribution of D_n is independent on $F_0(x)$, so the test is distribution-free.

More details on the classic Kolmogorov-Smirnov test for continuous distribution functions are given in Conover (1999), D'Agostino and Stephens (1986), Facchinetti (2009), Facchinetti (2011), Feller (1948), Frosini (1978), Gibbons and Chakraborti (2003), Kendall and Stuart (1967).

When the continuity assumption is not satisfied, D_n does not depends on the probability integral transformation of the sample values, and the probability distribution of D_n depends on $F_0(x)$, thus the test is not distribution-free.

2. SURVEY OF THE PREVIOUS WORKS

Several authors have studied the problem of the application of Kolmogorov-Smirnov test for discrete random variables.

Kolmogorov, Noether and Walsh showed that the Kolmogorov-Smirnov test is conservative when the random variable X is not continuous. In particular Kolmogorov (1941) first proved that whatever is the distribution function $F_0(x)$,

$$P(D_n \leq d) \geq H_n(d)$$

where $H_n(d)$ is the distribution function of D_n statistic for $F_0(x)$ continuous. Moreover, he established the confidence limit for $F_0(x)$ free from any restriction concerning the nature of the function.

Based on Kolmogorov studies, Noether (1963) provided a simplified proof by examining the connection between the distribution of D_n statistic and the distribution of D'_n statistic defined by

$$D'_n = \max_{x_i} |G_n(x_i) - G_0(x_i)|$$

where $G_n(x)$ and $G_0(x)$ are respectively the empirical and the theoretical distribution functions of a discrete random variable X.

Walsh (1963) proposed a method to show that the values for the continuous case represent bounds for the values that occur in any discrete case, based on the consideration that any discrete data situation can be interpreted as a situation involving the grouping of continuous data. According to this observation all data can be treated as discrete.

Schmid (1958) examined the asymptotic distributions of the Kolmogorov-Smirnov statistics when the hypothesized distribution function possesses a finite number of discontinuities. In particular, the author first extended the limit theorems of Kolmogorov-Smirnov and showed, by a generalization of the method of Kolmogorov (1933), that the probabilities converge also in this case. Furthermore, he observed that the limiting distribution depends on the values of $F_0(x)$ at the discontinuity points. Unfortunately this distribution function becomes undefined when $F(x)$ is purely discrete.

Also Carnal (1962) and Wood and Altavela (1978) like Schmid derived the limiting distribution of the statistic in the case of a discontinuous distribution function. Carnal (1962) considered a generalization of the method proposed by Feller (1948) instead that by Kolmogorov, and Wood and Altavela (1978) applied a result due to Billingsley (1968) on the weak convergence of the sample distribution function. This procedure requires only that the statistic can be rewritten as a continuous functional of a specific empirical process.

Conover (1972) and Coberley and Lewis (1972) gave the formulae in terms of recursive relationships for the exact distribution function and the power of the one-sided Kolmogorov-Smirnov statistics for discontinuous distributions. In particular, Conover presented a review of the use of the Kolmogorov-Smirnov statistic with discrete or discontinuous distribution, and provided an approximation to the distribution of the two-sided statistic. However, the same author observed that these computations are not feasible for large sample sizes and suggested to use the discrete Kolmogorov-Smirnov test for $n \leq 30$.

Pettitt and Stephens (1977) showed how the Kolmogorov-Smirnov statistic should be used for a test of fit for discrete or grouped data for sample sizes 30 or less. In particular, they evaluated the distribution function and the power of D_n by using Steck's or Durbin's formulae. The authors also made a comparison between Kolmogorov-Smirnov statistic for discrete data and Pearson Chi-square statistic.

Jalla and Marvulli (1957) presented a group of results based on an extensive numerical analysis of the properties of Kolmogorov-Smirnov test. In particular the authors evaluated the level of approximation of the test for discrete distributions. Afterwards Jalla (1979) gave a more general and complete form of the original proof, and Marvulli (1980) provided a set of operating tables for the use of the test in the discrete case.

Azzalini and Diana (1981) suggested a procedure to compute the distribution function of Kolmogorov-Smirnov test for discrete distribution with finite or countable number of jump points. In particular the problem for infinite jumps has been reduced to an equivalent one for finite jumps. The authors also applied the procedure when the null hypothesis assumes a Poisson distribution with specified parameter. For this case they evaluated the power of the test by using a Monte Carlo simulation and made a comparison with the power of the Pearson Chi-square test.

3. TEST PROCEDURE

Let X be a discrete random variable that assumes values in the finite set $S_x = \{x_i : i = 1, \ldots, m\}$ with cumulative distribution function:

$$F(x) = \begin{cases} 0 & \text{if } x < x_1 \\ \sum_{i=1}^{h} p_i & \text{if } x_h \leq x < x_{h+1} \\ 1 & \text{if } x \geq x_m. \end{cases} \quad \text{with} \quad h = 1, \ldots, m-1 \tag{4}$$

where $p_i = P(X = x_i)$. Since X is discrete, this is a function with m steps of height equal to the probabilities p_i for $i = 1, \ldots, m$.

Let $(x_{(1)} \leq \ldots \leq x_{(i)} \leq \ldots \leq x_{(n)})$ be an ordered random sample of size n from X with values $x_{(i)} \in S_x$ for $i = 1, \ldots, n$ not necessarily distinct.

The empirical cumulative distribution function is:

$$S_n(x) = \begin{cases} 0 & \text{if } x < x_1 \\ \sum_{i=1}^{h} \hat{p}_i & \text{if } x_h \leq x < x_{h+1} \\ 1 & \text{if } x \geq x_m. \end{cases} \quad \textit{with} \quad h = 1, \ldots, m-1 \tag{5}$$

with $\hat{p}_i = \frac{n_i}{n}$, and where n_i is the number of the observations $x_{(i)}$ in the sample equal to different values x_i, $n_i \in N^+$ and $\sum_{i=1}^{m} n_i = n$.

This is a step function with steps height equal to the estimated probabilities $\hat{p}_i$ $(i = 1, \ldots, m)$ i.e. equal to the multiple of $\frac{1}{n}$.

To compare the theoretical distribution function $F(x)$ with the empirical ones $S_n(x)$, we can refer to Figure 1.

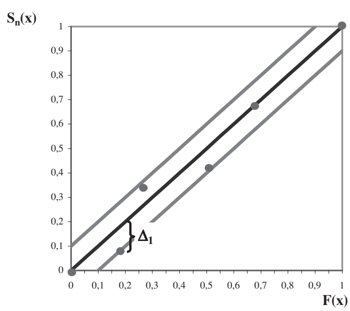

FIGURE 1. - *Comparison between $F(x)$ and $S_n(x)$*

The figure shows a trajectory of points that are aligned on the bisector if the two functions overlap, while escaping from it otherwise.

It also shows the differences Δ_h between the two functions:

$$\begin{array}{ll} \Delta_1 = S_1(x) - F_1(x) & \text{if } x_{(1)} \le x < x_{(2)} \\ \vdots & \vdots \\ \Delta_{m-1} = S_{m-1}(x) - F_{m-1}(x) & \text{if } x_{(m-1)} \le x < x_{(m)} \\ \Delta_m = S_m(x) - F_m(x) = 0 & \text{if } x \ge x_{(m)} \end{array} \tag{6}$$

as the vertical distances between the points and the bisector.

Consequently, the Kolmogorov-Smirnov statistic is simply the value of the largest absolute vertical difference between them, i.e.:

$$\Delta = \sup_{j=1,\dots,m-1} |\Delta_j|. \tag{7}$$

For a fixed value $0 \le D \le 1$, we define the two straight lines:

$$\begin{cases} y = x + D & \text{upper line } r_1; \\ y = x - D & \text{lower line } r_2 \end{cases}. \tag{8}$$

which are parallel to the bisector of Figure 1.

In order to apply the test and to define the critical values, we need to specify the cumulative distribution function of the test statistic under the null hypothesis

$$F_\Delta(D) = P\{\Delta \le D\}. \tag{9}$$

Once defined the distribution of the discrete random variable X, for a fixed value $F(x_{(h)}) = F_{(h)}$ $(h = 1, \dots, m-1)$ of the theoretic distribution function, we can check the following three situations:

1 $S_n(x) > F_{(h)} + D \Rightarrow$ the point falls above the upper line;

2 $F_{(h)} - D \le S_n(x) \le F_{(h)} + D \Rightarrow$ the point falls within the two lines;

3 $S_n(x) < F_{(h)} - D \Rightarrow$ the point falls below the lower line.

In particular, the probability (9) is determined by considering all possible trajectories that join the points $(F_{(h)}, S_h(x))$ for $h = 1, \dots, m$ that meet the condition 2. We must be aware that the possible events are only those that occur inside the unit square. As a consequent, the following conditions must be satisfied:

$$\begin{cases} F_{(h)} + D < 1 \\ F_{(h)} - D > 0 \end{cases}. \tag{10}$$

Let now h_M be the maximum value of h that satisfied the first condition of the system (10), and let h_m be the minimum value of h which meet the second condition of the system (10). It is obtained that the possible events are only for $h = 0, \dots, h_M$ with reference to the upper line and $h = h_m, \dots, m-1$ with reference to the lower one.

Thus we can accept all points $(F_{(h)}, S_h(x))$ for $h = 1, \ldots, m$ such that:

$$\begin{cases} S_h(x) \leq \frac{r_{1h}}{n} & \text{for } h = 0, \ldots, h_M \\ S_h(x) \geq \frac{r_{2h}}{n} & \text{for } h = h_m, \ldots, m \end{cases} \tag{11}$$

where

$$\frac{r_{1h}}{n} = \frac{int[n(F_{(h)} + D)]}{n} \quad \text{and} \quad \frac{r_{2h}}{n} = \frac{int[n(F_{(h)} - D)] + 1}{n}$$

are $\forall h$ the maximum and the minimum values of the empirical distribution function such that the points $(F_{(h)}, S_h(x))$ do not overrun the two lines in (8).

It is therefore possible to define the probability of occurrence of all possible trajectories by a multinomial probability mass function

$$P(n_1, \ldots, n_m) = \frac{n!}{\prod_{h=1}^{m} n_h!} \prod_{h=1}^{m} p_h^{n_h} \tag{12}$$

where p_h, for $h = 1, \ldots, m$, are the probabilities arising from the probability function of the discrete random variable X.

Since the different trajectories are all distinct, the probability of being inside the region defined by the two straight lines in (8) is:

$$Pr(\Delta \leq D) = \sum_{\{n_h : n_h \in I\}} P(n_1, \ldots, n_m) \tag{13}$$

where $I = \bigcup_{h=1}^{m-1} \{max(0, r_{1h}) \leq n_h \leq min(n, r_{2h})\}$.

4. DISTRIBUTION OF THE TEST STATISTIC FOR BINOMIAL RANDOM VARIABLES

The proposed procedure allows to define the exact critical values of Kolmogorov-Smirnov test under the null hypothesis, when X is a discrete random variable. The problem is that the number of different possible trajectories contained between the two straight lines grows up more than exponentially with increasing of the sample size n. Consequently in our analysis we considered sample sizes not exceeding $n = 20$. However, it is possible to extend the procedure to higher values of n.

Let X be a binomial random variable, we calculated the probability

$$F_\Delta(D^*) = Pr(\Delta \leq D^*) \tag{14}$$

where the reference values D^* are those corresponding to the continuous case for the significance level $\alpha = 0.1$ (see Massey, 1951; Birnbaum, 1952; Facchinetti, 2011). In particular, for $X \sim Bin(k;p)$ with $p = 0.5$, $k = 1, 2, \ldots, 10$ and sample sizes $n = 5, 10, 15, 20$, the values of the distribution function of the Δ statistic are in Table 1.

TABLE 1. - $F_{\Delta}(D^*)$ *for a binomial random variable with* $p = 0.5$*, and* $\alpha = 0.1$

	n	5	10	15	20
	D^*	**0.50945**	**0.36866**	**0.30397**	**0.26473**
	1	1.0000	0.9785	0.9926	0.9882
	2	0.9688	0.9930	0.9916	0.9921
	3	0.9978	0.9707	0.9904	0.9848
	4	0.9899	0.9728	0.9829	0.9840
k	**5**	0.9891	0.9709	0.9876	0.9855
	6	0.9894	0.9873	0.9700	0.9800
	7	0.9755	0.9581	0.9800	0.9760
	8	0.9849	0.9859	0.9852	0.9791
	9	0.9588	0.9717	0.9825	0.9778
	10	0.9812	0.9652	0.9731	0.9625

If we now consider the binomial random variable $X \sim Bin(k; p)$ with $p = 0.1$, $k = 1, 2, \ldots, 10$ and sample sizes $n = 5, 10, 15, 20$, the values of the distribution function of the Δ statistic are in Table 2.

TABLE 2. - $F_{\Delta}(D^*)$ *for a binomial random variable with* $p = 0.1$*, and* $\alpha = 0.1$

	n	5	10	15	20
	D^*	**0.50945**	**0.36866**	**0.30397**	**0.26473**
	1	0.9995	0.9984	0.9997	0.9996
	2	0.9945	0.9951	0.9970	0.9983
	3	0.9788	0.9942	0.9925	0.9905
	4	0.9940	0.9957	0.9873	0.9924
k	**5**	0.9844	0.9799	0.9837	0.9883
	6	0.9769	0.9892	0.9827	0.9765
	7	0.9726	0.9763	0.9793	0.9761
	8	0.9797	0.9700	0.9948	0.9770
	9	0.9798	0.9719	0.9778	0.9844
	10	0.9738	0.9902	0.9800	0.9997

To verify the validity of the obtained results, we have made a Monte Carlo simulation. In particular, in Tables 3 and 4 the simulated values of the distribution function of the Δ statistic, indicated with $F_\Delta(\tilde{D})$, for the binomial random variable $X \sim Bin(k; p)$ considered above are presented.

TABLE 3. - *$F_\Delta(\tilde{D})$ simulated for a binomial random variable with $p = 0.5$, and $\alpha = 0.1$*

	n	**5**	**10**	**15**	**20**
	D^*	**0.50945**	**0.36866**	**0.30397**	**0.26473**
	1	1.0000	0.9787	0.9930	0.9963
	2	0.9627	0.9923	0.9947	0.9980
	3	0.9987	0.9703	0.9877	0.9813
	4	0.9877	0.9680	0.9790	0.9843
k	**5**	0.9890	0.9717	0.9890	0.9857
	6	0.9893	0.9853	0.9793	0.9867
	7	0.9757	0.9563	0.9727	0.9737
	8	0.9810	0.9887	0.9863	0.9773
	9	0.9600	0.9723	0.9827	0.9770
	10	0.9747	0.9680	0.9723	0.9623

TABLE 4. - *$F_\Delta(\tilde{D})$ simulated for a binomial random variable with $p = 0.1$, and $\alpha = 0.1$*

	n	**5**	**10**	**15**	**20**
	D^*	**0.50945**	**0.36866**	**0.30397**	**0.26473**
	1	1.0000	0.9987	0.9997	0.9993
	2	0.9957	0.9963	0.9953	0.9970
	3	0.9753	0.9920	0.9930	0.9910
	4	0.9953	0.9950	0.9903	0.9917
k	**5**	0.9787	0.9797	0.9893	0.9877
	6	0.9790	0.9867	0.9813	0.9843
	7	0.9743	0.9703	0.9803	0.9750
	8	0.9783	0.9700	0.9917	0.9817
	9	0.9823	0.9723	0.9807	0.9787
	10	0.9777	0.9920	0.9807	0.9900

From Tables 1-4 we observe the closeness of the values provided by the proposed procedure with those simulated.

To allow a synthetic comparison between exact and simulated critical values in Tables 2 and 4 for $X \sim Bin(k; 0.1)$, Table 5 gives the percentage differences based on the simulated values

$$\frac{F_{\Delta}(D^{*}) - F_{\Delta}(\tilde{D})}{F_{\Delta}(\tilde{D})}.$$

TABLE 5. - *Percentage differences between critical values in Tables 2 and 4*

	n	**5**	**10**	**15**	**20**
	D^*	**0.50945**	**0.36866**	**0.30397**	**0.26473**
	1	-0.0500%	-0.0267%	0.0033%	0.0267%
	2	-0.1172%	-0.1238%	0.1674%	0.1304%
	3	0.3554%	0.2218%	-0.0504%	-0.0505%
	4	-0.1340%	0.0704%	-0.3063%	0.0739%
k	**5**	0.5858%	0.0238%	-0.5694%	0.0641%
	6	-0.2145%	0.2568%	0.1393%	-0.7958%
	7	-0.1779%	0.6149%	-0.1054%	0.1128%
	8	0.1397%	0.0000%	0.3160%	-0.4754%
	9	-0.2579%	-0.0446%	-0.2923%	0.5858%
	10	-0.3955%	-0.1825%	-0.0680%	0.9798%

From Table 5 we observe that the minimum and the maximum percentage differences are respectively -0.7958% and 0.9798%. Being these values less of 1 percentage point, we confirm that the proposed procedure can be used to calculate the exact critical values of the Kolmogorov-Smirnov test in the discrete case.

Furthermore, fixed the significance level α, it is possible to make a comparison, in terms of probability, between the exact critical values for the discrete and the continuous case, considering the critical values tabulated by Massey (1951) and then integrated by Birnbaum (1952).

In particular, the following figures show, for $n = 5$, the distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases to give an idea of the trend of these functions.

We observe that for the continuous case the distribution function is always the same, while in the discrete case that function depends to the parameters of the binomial random variable considered.

Looking at the Figures 2 and 3 we observe that there is a confirmation of what proved by Kolmogorov (1941), Noether (1963) and Walsh (1963): if X is a discrete random variable, the critical values are lower or equal to the corresponding critical values of the continuous case, i.e. the test is conservative. This highlights a more precautionary procedure for the discrete case respect to the continuous ones.

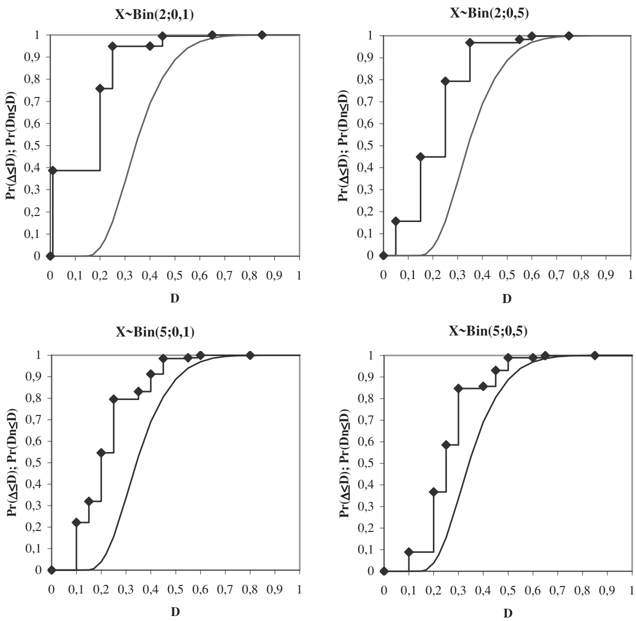

FIGURE 2. - *Comparison of the distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases;* $n = 5$

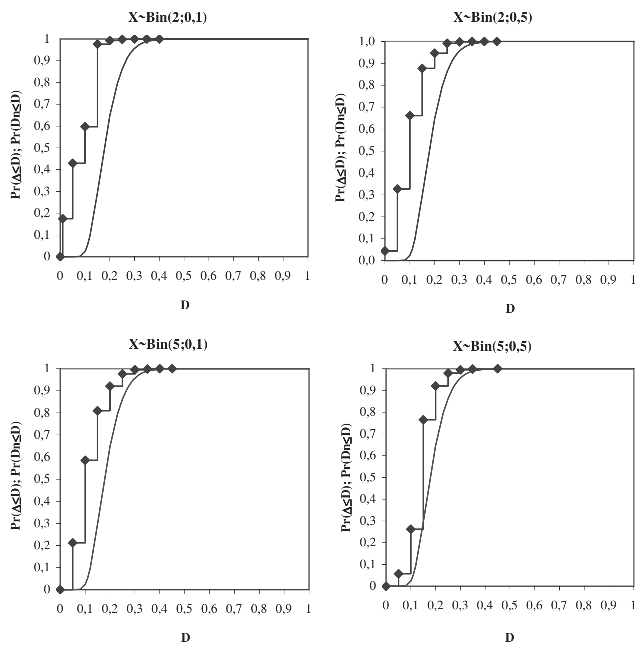

FIGURE 3. - *Comparison of the distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases;* $n = 20$

We also observe that for increasing k and for $p \to 0.5$, the differences between continuous and discrete cases tend to weaken. In addition, improving the sample size from $n = 5$ to $n = 20$, we note a lower range of variation between the curved line and the piecewise ones.

5. DISTRIBUTION OF THE TEST STATISTIC FOR THREE EXTREME DISCRETE RANDOM VARIABLES

We now consider three particular discrete random variables: the uniform, the triangle negative and the triangle symmetric, all with parameter $m = 5$, as showed in Figure 4:

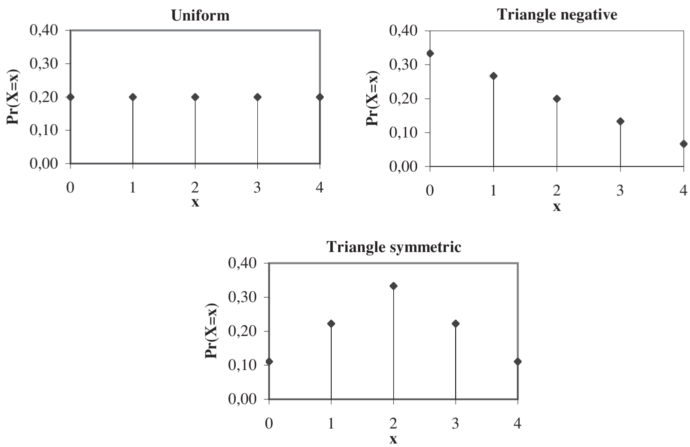

FIGURE 4. - *Probability distribution function of the three extreme discrete random variables*

Fixed the significance level $\alpha = 0.1$ and referring to the value $D^* = 0.50945$ obtained in the continuous case for sample size $n = 5$, we calculated the values of the probability $F_\Delta(D^*)$ for the three examined distributions.

The obtained values are:

1 uniform distribution: $Pr(\Delta \leq D^*) = 0.96992$;
2 triangle negative distribution: $Pr(\Delta \leq D^*) = 0.97907$;
3 triangle symmetric distribution: $Pr(\Delta \leq D^*) = 0.99075$.

We note that also in these cases the probabilities related to the discrete case are higher than the fixed value $1 - \alpha = 0.9$ of the corresponding continuous case.

The following figures represent, for $n = 5$, the cumulative distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases to give an idea of the trend of these functions.

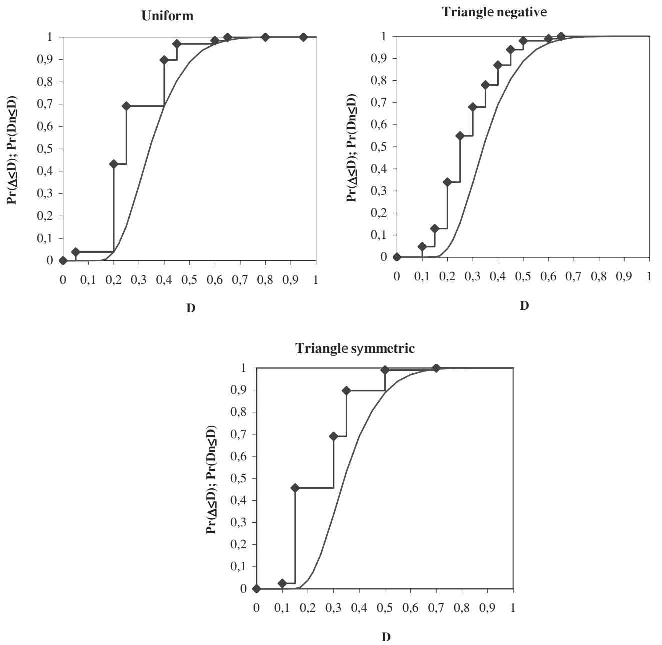

FIGURE 5. - *Comparison of the distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases; $n = 5$*

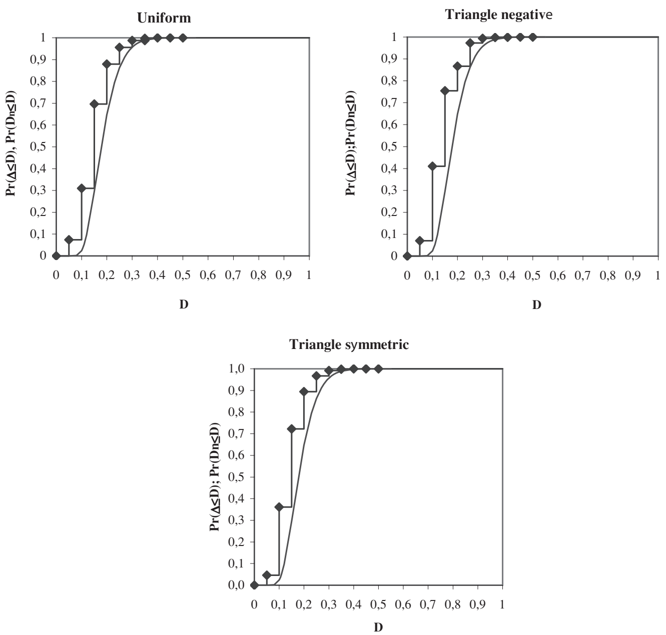

FIGURE 6. - *Comparison of the distribution function of Kolmogorov-Smirnov statistic for continuous and discrete cases; $n = 20$*

In Figures 2, 3, 5 and 6 we observe a similar trend. Further, in Figures 5 and 6 we observe that the values of the three extreme discrete distributions are closer to those calculated for continuous case respect to those obtained for the binomial distributions. In particular, in these cases, the distribution function $Pr(\Delta \leq D^*)$ is almost coincident with that of the continuous case.

In addition, increasing the sample size from $n = 5$ to $n = 20$, there is also in this case a lower range of variation between the curved line and the piecewise ones.

6. CONCLUSIONS

The Kolmogorov-Smirnov test is one of the most important goodness of fit tests.

In this paper we have adapted such test, valid for continuous random variables, to evaluate the goodness of fit of a specified discrete random variable.

The proposed statistic is

$$\Delta = \sup_{j=1,\dots,m-1} |S_j(x) - F_j(x)|. \tag{15}$$

The procedure is based on the calculus of a multinomial distribution function that depends on the sample size n, on the distribution law of the discrete random variable X and on its parameters. Consequently, in this case the statistic is not distribution-free with respect to the classic Kolmogorov-Smirnov statistic for a continuous random variable.

These observations are confirmed by the conducted applications for binomial, uniform, triangle negative and triangle symmetric distributions.

Nevertheless, the proposed procedure can be effectively used to verify the null hypothesis $H_0 : F(x) = F_0(x)\ \forall x$ against the alternative $H_1 : \exists x : F(x) \neq F_0(x)$ with $F_0(x)$ known. In this case, it is possible to obtain, at the same significance level α, a wider region to refuse the null hypotheses H_0, and therefore a greater power of the test, compared to the one obtained using the critical values provided by the classic Kolmogorov-Smirnov test.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. U. Magagnoli for the supervision of the work.

REFERENCES

Azzalini A., Diana G. (1981). The distribution of Kolmogorov's statistic for discrete variates. *Metron*, **XXXIX**(3-4), 133-140.

Billingsley P. (1968). *Convergence of probability measures*. John Wiley & Sons, New York.

Birnbaum Z.W. (1952). Numerical tabulation of the distribution of Kolmogorov statistic for finite sample size. *Journal of the American Statistical Association*, **47**, 425-441.

Carnal H. (1962). Sur les theoremes de Kolmogorov et Smirnov dans le cas d'une distribution discontinue. *Commentarii Mathematici Helvetici*, **37**(1), 19-35.

Coberley W.A., Lewis T.O. (1972). A note on a one-sided Kolmogorov-Smirnov test of fit for discrete distribution functions. *The Annals of Institute of Statistical Mathematics*, **24**(1), 183-187.

Conover W.J. (1972). A Kolmogorov goodness of fit test for discontinuous distributions. *Journal of the American Statistical Association*, **67**(339), 591-596.

Conover W.J. (1999). *Practical nonparametric statistic*. John Wiley & Sons, New York.

D'Agostino R.B., Stephens M.A. (1986). *Goodness of fit techniques*. Marcell Dekker, New York.

Facchinetti S. (2009). Una procedura per determinare i valori critici esatti per il test di Kolmogorov-Smirnov. *Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore*, **Serie E.P. 135**, 1-24.

Facchinetti S. (2011). A procedure to find exact critical values of Kolmogorov-Smirnov test. *Statistica Applicata*, accepted for publication.

Feller W. (1948). On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Annals of Mathematical Statistics*, **19**(2), 177-189.

Frosini B.V. (1978). A survey of class of goodness of fit statistics. *Metron*, **XXXVI**(1-2), 3-49.

Gibbons J.D., Chakraborti S. (2003). *Nonparametric statistical inference*. Marcel Dekker, New York.

Jalla E. (1979). Il test di Kolmogorov nel caso di distribuzione discreta. *Istituto di Statistica dell'Università degli Studi di Torino*, **4**, 1-16.

Jalla E., Marvulli R. (1957). Sul test di Kolmogorov nel caso di distribuzione discreta. *Atti della XXVIII riunione scientifica della Società Italiana di Statistica*, Padova, **2**.

Kendall M.G., Stuart A. (1967). *The advanced theory of statistics*. Griffin, London.

Kolmogorov A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, **4**, 83-91.

Kolmogorov A. (1941). Confidence limits for an unknown distribution function. *The Annals of Mathematical Statistics*, **12**(4), 461-463.

Marvulli R. (1980). Tabelle per l'uso del test di Kolmogorov nel caso discreto. *Istituto di Statistica dell'Università degli Studi di Torino*, **6**.

Massey F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, **46**, 68-78.

Noether G.E. (1963). Note on the Kolmogorov statistic in the discrete case. *Metrika*, **7**(2), 115-116.

Pettitt A.N., Stephens M.A. (1977). The Kolmogorov-Smirnov goodness of fit statistic with discrete and grouped data. *Journal of the American Statistical Association*, **19**(2), 205-210.

Schmid P. (1958). On the Kolmogorov and Smirnov limit theorems for discontinuous distribution functions. *The Annals of Mathematical Statistics*, **29**(4), 1011-1027.

Smirnov N. (1939). Sur les ecarts de la coubre de distribution empirique. *Recueil Mathematique*, **6**, 3-26.

Walsh J.E. (1963). Bounded probability properties of Kolmogorov-Smirnov and similar statistics for discrete data. *The Annals of Institute of Statistical Mathematics*, **15**(2), 153-158.

Wood C.L., Altavela M.M. (1978). Large-sample results for Kolmogorov-Smirnov test for discrete distributions. *Biometrika*, **65**(1), 235-239.