Journal of Econometrics 162 (2011) 326–344

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

# Generalized runs tests for the IID hypothesis

Jin Seo Cho [a], Halbert White [b,*]

[a] School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, South Korea

[b] Department of Economics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

ARTICLE INFO

*Article history:*
Received 5 August 2009
Received in revised form
27 January 2011
Accepted 3 February 2011
Available online 12 February 2011

*JEL classification:*
C12
C23
C80

*Keywords:*
IID condition
Runs test
Geometric distribution
Gaussian process
Dependence
Structural break

ABSTRACT

We provide a family of tests for the IID hypothesis based on generalized runs, powerful against unspecified alternatives, providing a useful complement to tests designed for specific alternatives, such as serial correlation, GARCH, or structural breaks. Our tests have appealing computational simplicity in that they do not require kernel density estimation, with the associated challenge of bandwidth selection. Simulations show levels close to nominal asymptotic levels. Our tests have power against both dependent and heterogeneous alternatives, as both theory and simulations demonstrate.



## 1. Introduction

The assumption that data are independent and identically distributed (IID) plays a central role in the analysis of economic data. In cross-section settings, the IID assumption holds under pure random sampling. As Heckman (2001) notes, violation of the IID property, therefore random sampling, can indicate the presence of sample selection bias. The IID assumption is also important in time-series settings, as processes driving time series of interest are often assumed to be IID. Moreover, transformations of certain time series can be shown to be IID under specific null hypotheses. For example Diebold et al. (1998) show that to test density forecast optimality, one can test whether the series of probability integral transforms of the forecast errors are IID uniform ($U[0, 1]$).

There is a large number of tests designed to test the IID assumption against specific alternatives, such as structural breaks, serial correlation, or autoregressive conditional heteroskedasticity. Such special purpose tests may lack power in other directions, however, so it is useful to have available broader diagnostics that may alert researchers to otherwise unsuspected properties of their data. Thus, as a complement to special purpose tests, we consider tests for the IID hypothesis that are sensitive to general alternatives. Here we exploit runs statistics to obtain necessary and sufficient conditions for data to be IID. In particular, we show that if the underlying data are IID, then suitably defined runs are IID with the geometric distribution. By testing whether the runs have the requisite geometric distribution, we obtain a new family of tests, the *generalized runs tests*, suitable for testing the IID property. An appealing aspect of our tests is their computational convenience relative to other tests sensitive to general alternatives to IID. For example, Hong and White's (2005) entropy-based IID tests require kernel density estimation, with its associated challenge of bandwidth selection. Our tests do not require kernel estimation and, as we show, have power against dependent alternatives. Our tests also have power against structural break alternatives, without exhibiting the non-monotonicities apparent in certain tests based on kernel estimators (Crainiceanu and Vogelsang, 2007; Deng and Perron, 2008).

Runs have formed an effective means for understanding data properties since the early 1940s. Wald and Wolfowitz (1940), Mood (1940), Dodd (1942) and Goodman (1958) first studied runs to test for randomness of data with a fixed percentile $p$ used in defining the runs. Granger (1963) and Dufour (1981) propose using runs as a nonparametric diagnostic for serial correlation, noting

* Corresponding author. Tel.: +1 858 534 5985; fax: +1 858 534 7040.
*E-mail addresses:* jinseocho@yonsei.ac.kr (J.S. Cho), hwhite@weber.ucsd.edu, drhalwhite@yahoo.com (H. White).

0304-4076/$ – see front matter 
doi:10.1016/j.jeconom.2011.02.001

that the choice of $p$ is important for the power of the test. Fama (1965) extensively exploits a runs test to examine stylized facts of asset returns in US industries, with a particular focus on testing for serial correlation of asset returns. Heckman (2001) observes that runs tests can be exploited to detect sample selection bias in cross-sectional data; such biases can be understood to arise from a form of structural break in the underlying distributions.

Earlier runs tests compared the mean or other moments of the runs to those of the geometric distribution for fixed $p$, say 0.5 (in which case the associated runs can be computed alternatively using the median instead of the mean). Here we develop runs tests based on the probability generating function (PGF) of the geometric distribution. Previously, Kocherlakota and Kocherlakota (KK, 1986) have used the PGF to devise tests for discrete random variables having a given distribution under the null hypothesis. Using fixed values of the PGF parameter $s$, KK develop tests for the Poisson, Pascal–Poisson, bivariate Poisson, or bivariate Neyman type A distributions. More recently, Rueda et al. (1991) study PGF-based tests for the Poisson null hypothesis, constructing test statistics as functionals of stochastic processes indexed by the PGF parameter $s$. Here we develop PGF-based tests for the geometric distribution with parameter $p$, applied to the runs for a sample of continuously distributed random variables.

We construct our test statistics as functionals of stochastic processes indexed by both the runs percentile $p$ and the PGF parameter $s$. By not restricting ourselves to fixed values for $p$ and/or $s$, we create the opportunity to construct tests with superior power. Further, we obtain weak limits for our statistics in situations where the distribution of the raw data from which the runs are constructed may or may not be known and where there may or may not be estimated parameters. As pointed out by Darling (1955), Sukhatme (1972), Durbin (1973) and Henze (1996), among others, goodness-of-fit (GOF) based statistics such as ours may have limiting distributions affected by parameter estimation. As we show, however, our test statistics have asymptotic null distributions that are not affected by parameter estimation under mild conditions. We also provide straightforward simulation methods to consistently estimate asymptotic critical values for our test statistics.

We analyze the asymptotic local power of our tests, and we conduct Monte Carlo experiments to explore the properties of our tests in settings relevant for economic applications. In studying power, we give particular attention to dependent alternatives and to alternatives containing an unknown number of structural breaks. To analyze the asymptotic local power of our tests against dependent alternatives, we assume a first-order Markov process converging to an IID process in probability at the rate $n^{-1/2}$, where $n$ is the sample size, and we find that our tests have nontrivial local power. We work with first-order Markov processes for conciseness. Our results generalize to higher-order Markov processes, but that analysis is sufficiently involved that we leave it for subsequent work.

Our Monte Carlo experiments corroborate our theoretical results and also show that our tests exhibit useful finite sample behavior. For dependent alternatives, we compare our generalized runs tests to the entropy-based tests of Robinson (1991), Skaug and Tjøstheim (1996) and Hong and White (2005). Our tests perform respectably, showing good level behavior and useful, and in some cases superior, power against dependent alternatives. For structural break alternatives, we compare our generalized runs tests to Feller's (1951) and Kuan and Hornik's (1995) RR test, Brown et al. 's (1975) RE-CUSUM test, Sen's (1980) and Ploberger et al. 's (1989) RE test, Ploberger and Krämer's (1992) OLS-CUSUM test, Andrews' (1993) Sup-$W$ test, Andrews and Ploberger's (1994) Exp-$W$ and Avg-$W$ tests, and Bai's (1996) $M$-test. These prior tests are all designed to detect a finite number of structural breaks at unknown locations. We find good level behavior for our tests and superior power against multiple breaks. An innovation is that we consider alternatives where the number of breaks grows with sample size. Our new tests perform well against such structural break alternatives, whereas the prior tests do not.

This paper is organized as follows. In Section 2, we introduce our new family of generalized runs statistics and derive their asymptotic null distributions. These involve Gaussian stochastic processes. Section 3 provides methods for consistently estimating critical values for the test statistics of Section 2. This permits us to compute valid asymptotic critical values even when the associated Gaussian processes are transformed by continuous mappings designed to yield particular test statistics of interest. We achieve this using other easily simulated Gaussian processes whose distributions are identical to those of Section 2. Section 4 studies aspects of local power for our tests. Section 5 contains Monte Carlo simulations; this also illustrates the use of the simulation methods developed for obtaining the asymptotic critical values in Section 2. Section 6 contains concluding remarks. All mathematical proofs are collected in the Appendix.

Before proceeding, we introduce mathematical notation used throughout. We let $\mathbf{1}_{\{\cdot\}}$ stand for the indicator function such that $\mathbf{1}_{\{A\}} = 1$ if the event $A$ is true, and 0 otherwise. $\Rightarrow$ and $\to$ denote 'converge(s) weakly' and 'converge(s) to', respectively, and $\overset{d}{=}$ denotes equality in distribution. Further, $\|\cdot\|$ and $\|\cdot\|_\infty$ denote the Euclidean and uniform metrics, respectively. We let $\mathcal{C}(A)$ and $\mathcal{D}(A)$ be the spaces of continuous and cadlag mappings from a set $A$ to $\mathbb{R}$, respectively, and we endow these spaces with Billingsley's (1968, 1999) or Bickel and Wichura's (1971) metric. We denote the unit interval as $\mathbb{I} := [0, 1]$.

## 2. Testing the IID hypothesis

### 2.1. Maintained assumptions

We begin by collecting together assumptions maintained throughout and proceed with our discussion based on these. We first specify the data generating process (DGP) and a parameterized function whose behavior is of interest.

A1 (DGP): *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a complete probability space. For* $m \in \mathbb{N}$, $\{\mathbf{X}_t : \Omega \mapsto \mathbb{R}^m,\ t = 1, 2, \ldots\}$ *is a stochastic process on* $(\Omega, \mathcal{F}, \mathbb{P})$.

A2 (Parameterization): *For* $d \in \mathbb{N}$, *let* $\boldsymbol{\Theta}$ *be a non-empty convex compact subset of* $\mathbb{R}^d$. *Let* $h : \mathbb{R}^m \times \boldsymbol{\Theta} \mapsto \mathbb{R}$ *be a function such that* (i) *for each* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $h(\mathbf{X}_t(\cdot), \boldsymbol{\theta})$ *is measurable; and* (ii) *for each* $\omega \in \Omega$, $h(\mathbf{X}_t(\omega), \cdot)$ *is such that for each* $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Theta}$, $|h(\mathbf{X}_t(\omega), \boldsymbol{\theta}) - h(\mathbf{X}_t(\omega), \boldsymbol{\theta}')| \le M_t(\omega)\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|$, *where* $M_t$ *is measurable and is* $O_{\mathbb{P}}(1)$, *uniformly in* $t$.

Assumption A2 specifies that $\mathbf{X}_t$ is transformed via $h$. The Lipschitz condition of A2(ii) is mild and typically holds in applications involving estimation. Our next assumption restricts attention to continuously distributed random variables.

A3 (Continuous random variables): *For given* $\boldsymbol{\theta}_* \in \boldsymbol{\Theta}$, *the random variables* $Y_t := h(\mathbf{X}_t, \boldsymbol{\theta}_*)$ *have continuous cumulative distribution functions (CDFs)* $F_t : \mathbb{R} \mapsto \mathbb{I},\ t = 1, 2, \ldots$.

Our main interest attaches to distinguishing the following hypotheses:

$$\mathbb{H}_0 : \{Y_t : t = 1, 2, \ldots\} \text{ is an IID sequence;}$$
$$\text{vs.} \quad \mathbb{H}_1 : \{Y_t : t = 1, 2, \ldots\} \text{ is not an IID sequence.}$$

Under $\mathbb{H}_0$, $F_t \equiv F$ (say), $t = 1, 2, \ldots$. We separately treat the cases in which $F$ is known or unknown. In the latter case, we estimate $F$ using the empirical distribution function.