Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Asymptotically Refined Score and GOF Tests for Inverse Gaussian Models

Anthony F. Desmond[a] and Zhenlin Yang[b]*

[a]*Department of Mathematics and Statistics, University of Guelph*

[b]*School of Economics, Singapore Management University*

emails: tdesmond@uoguelph.ca, zlyang@smu.edu.sg

January 12, 2016

Abstract

The score test and the GOF test for the inverse Gaussian distribution, in particular the latter, are known to have large size distortion and hence unreliable power when referring to the asymptotic critical values. We show in this paper that with the appropriately bootstrapped critical values, these tests become second-order accurate, with size distortion being essentially eliminated and power more reliable. Two major generalizations of the score test are made: one is to allow the data to be right-censored, and the other is to allow the existence of covariate effects. A data mapping method is introduced for bootstrap to be able to produce censored data that are conformable with the null model. Monte Carlo results clearly favour the proposed bootstrap tests. Real data illustrations are given.

Key Words: Bootstrap critical value; Data Mapping; Goodness of fit; Score test; Inverse Gaussian regression; Right-censoring; Wiener process.

AMS subject classification: 62F03, 62F05, 62F40, 62N01,

1 Introduction

The inverse Gaussian distribution has become a popular model in a variety of application areas in the past few decades (Lancaster, 1972; Banerjee and Bhattacharyya, 1976; Chhikara and Folks, 1977, 1978, 1989; Whitmore, 1979, 1986; Edgeman, 1989, 1990; Pavur et al., 1992; Seshadri, 1993, 1998; Desmond and Chapman, 1993; Johnson et al., 1994; Hawkins and Olwell, 1997; Ducharme, 2001; Lee and Whitmore, 2006, 2010; Desmond and Yang, 2011). The suitability of the inverse Gaussian model can be assessed via goodness of fit (GOF) tests. This has been studied previously based on the empirical distribution function (EDF) (Edgeman, 1990; Pavur et al., 1992), a graphical method with standardized recursive residuals (Letac et al., 1985), GOF tests (Ducharme, 2001), and score tests (Desmond and Yang, 2011).

Our motivation is based on the underlying Wiener process model. Even with the inclusion of covariates in the drift parameter, there may be unexplained heterogeneity, which varies from

*Corresponding to: Zhenlin Yang, School of Economics, Singapore Management University, Singapore 178903; Telephone: 65-6828-0852; Fax: 65-6828-0833.

individual to individual. This is analogous to a frailty effect in more conventional survival models. Testing for the existence of frailty effects is an important topic in conventional models; this motivates our emphasis on testing versus IG models with random drift. Our motivation for using the refined bootstrap methods is that first order tests are known to be very poor in this case. In particular, conventional score tests need very large samples in order to be at all accurate in terms of level. The very unique feature of the score test, i.e., requiring only the estimation of the null model, makes it practically feasible to bootstrap the finite sample critical values, resulting in second-order accurate score tests. Also another important motivation is to extend goodness of fit tests to IG regression and censored data situations. Neither of these aspects has been treated, at all, previously in the literature, with the exception of Ducharme (2001), who has a short section on the censored situation but without covariates.

In this paper, we present asymptotically refined score and GOF tests of an inverse Gaussian distribution. We first show that the score test of Desmond and Yang (2011) and the GOF test of Ducharme (2001) have a second-order accuracy in rejection probability when referring to the appropriately bootstrapped critical values. Then, two major generalizations of the score test are made: one is to allow data to be right-censored, and the other is to allow the existence of covariate effects. To generate censored bootstrap samples that mimic the samples from the null model, a data mapping method is introduced. The same types of generalizations do not seem to be as simple for the other type of tests such as Ducharme's GOF test. Extensive Monte Carlo results clearly favour the proposed bootstrap tests.

Section 2 gives a complete treatment of the score and GOF tests for inverse Gaussian distribution based on a complete sample. Section 3 generalizes the score test to the case of a censored sample along with its bootstrap version. Section 4 further generalizes the score and bootstrap score tests to the case of an inverse Gaussian regression model, with complete or censored data. Section 5 presents Monte Carlo results. Section 6 provides some real data illustrations. Section 7 concludes the paper.

2 Score and GOF Tests based on a Complete Sample

2.1 Inverse Gaussian and inverse Gaussian mixtures

It is well known that the inverse Gaussian (`IG`) distribution arises as the *first hitting time* T of a Wiener process (with a *drift* coefficient μ, *diffusion* coefficient σ, starting at position 'zero') to a *boundary* or *barrier* ω. It has the following probability density function (pdf):

$$f(t|\omega,\mu,\sigma^2) = \frac{\omega}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{(\mu t-\omega)^2}{2\sigma^2 t}\right),\ t \geq 0, \tag{2.1}$$

where $\omega > 0$, $\mu > 0$, and $\sigma^2 > 0$. See, e.g., Chhikara and Folks (1989) and Aalen et al. (2008). Its cumulative distribution function (CDF) takes the form:

$$F(t|\omega, \mu, \sigma^2) = \Phi\left(\frac{\mu t - \omega}{\sqrt{\sigma^2 t}}\right) + \exp\left(\frac{2\omega\mu}{\sigma^2}\right)\Phi\left(-\frac{\mu t + \omega}{\sqrt{\sigma^2 t}}\right), \tag{2.2}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. Clearly, the pdf (2.1) is proper when $\mu > 0$ in the sense that it integrates to one or the limit of CDF (2.2) is 1 as $t \to \infty$. When $\mu < 0$, (2.1) is not a proper pdf in the sense that it does not integrate to one. When $\mu = 0$, the distribution is proper, i.e. integrates to 1. However, it is not of inverse Gaussian form and the mean does not exist (i.e. is infinite). It is a stable distribution of index .5; see Feller (1966, Vol.2, p.170). We will consider in this paper only the proper `IG` distribution.

There are three parameters in (2.1) but only two of them are free, which can easily be seen through reparameterization: $\delta = \mu/\omega$ and $\lambda = \omega^2/\sigma^2$, which is in fact a more popular form of parameterization and is followed in Desmond and Yang (2011):

$$f(t|\delta, \lambda) = \left(\frac{\lambda}{2\pi t^3}\right)^{1/2} \exp\left(-\frac{\lambda(\delta t - 1)^2}{2t}\right), \; t \geq 0, \; \delta \geq 0, \; \lambda \geq 0. \tag{2.3}$$

To remove the redundant parameter in (2.1), we follow the arguments of Lee and Whitmore (2006) (the underlying stochastic process used to describe an individual's 'health status' is usually latent, and an arbitrary unit can be used to measure such a process) and set $\sigma = 1$. Both parameterizations have merits: the former, $\texttt{IG}(\mu, \omega)$, allows the covariate effects to be incorporated more naturally through the drift parameter μ as well as the barrier ω, whereas the latter, $\texttt{IG}(\delta, \lambda)$, gives a simpler expression for the maximum likelihood estimators (MLE). In this paper, we will mainly use the (μ, ω) parameterization.

The pure `IG` distribution described above assumes that the individuals are homogeneous in the sense that they 'drift' in the same manner to the boundary. The *drift* may vary from individual to individual as some degree of "tracking" can often take place. In this case, it is typical to treat μ as a random variable, e.g., $\mu \sim N(m, \upsilon)$. Now (2.1) is viewed as the conditional pdf of T given μ. Integrating the joint pdf of T and μ with respect to the pdf of μ leads to the 'marginal' distribution of T as

$$g(t|\omega, m, \upsilon) = \frac{\omega}{\sqrt{2\pi}} \frac{1}{\sqrt{t^3(t\upsilon + 1)}} \exp\left(-\frac{(mt - \omega)^2}{2t(t\upsilon + 1)}\right), \tag{2.4}$$

where $\omega > 0$ and $\upsilon \geq 0$. This distribution is often referred to as the inverse Gaussian mixture (`IGM`) in the literature (Whitmore, 1986, Aalen, 1994), which offers a natural way to test the goodness of fit of the pure inverse Gaussian models. Clearly, it is a defective pdf as the integration w.r.t. μ is done over the entire real line, where whenever μ is negative, the conditional pdf of T given μ is an improper pdf. This can be seen from its CDF as well:

$$G(t|\omega, m, \upsilon) = \Phi\left(\frac{mt - \omega}{\sqrt{t^2\upsilon + t}}\right) + \exp(2\omega m + 2\omega^2\upsilon)\ \Phi\left(-\frac{\omega + 2\omega\upsilon t + mt}{\sqrt{t^2\upsilon + t}}\right), \tag{2.5}$$

(see, e.g, Aalen, 1994). From (2.5) above, it is easy to see that

$$\pi(\omega, m, \upsilon) = \lim_{t\to\infty} G(t|\omega, m, \upsilon) = \Phi\left(\frac{m}{\sqrt{\upsilon}}\right) + \exp(2\omega m + 2\omega^2\upsilon)\ \Phi\left(-\frac{2\omega\upsilon + m}{\sqrt{\upsilon}}\right),$$

which is less than 1 when $\upsilon > 0$ as it corresponds to the inverse Gaussian CDF evaluated at υ^{-1} (see Whitmore, 1986). Clearly, as $\upsilon \to 0$, $\pi(\omega, m, \upsilon) \to 1$ and $G(t|\omega, m, \upsilon) \to F(t|\omega, m)$.

While the defective feature of (2.4) may not affect the null behavior of the tests for an IG distribution, to implement the bootstrap procedure, a proper mixture distribution is needed (see the discussions later). This can be obtained in several ways, of which the simplest may be through rescaling:

$$h_1(t|\omega, m, \upsilon) = g(t|\omega, m, \upsilon)/\pi(\omega, m, \upsilon). \tag{2.6}$$

Desmond and Yang (2011) considered a truncated normal mixing distribution to restrict μ to be positive so that the resulting IGM is a proper distribution:

$$h_2(t|\omega, m, \upsilon) = g(t|\omega, m, \upsilon)\Phi\left(\frac{(\upsilon + m)\omega}{\sqrt{\upsilon^2 t + \upsilon}}\right)\Phi^{-1}\left(\frac{m\omega}{\sqrt{\upsilon}}\right), \tag{2.7}$$

and they further proposed a general way of handling the problems with the defective mixtures by allowing the mixing distribution to be an arbitrary distribution with mean m and variance υ. The resulting marginal distribution of T is approximately

$$h_3(t|\omega, m, \upsilon) = f(t|\omega, m)\big[1 + \frac{1}{2}\upsilon(\omega^2(mt-1)^2 - t)\big], \tag{2.8}$$

which is shown to be a proper pdf if $m > \frac{\upsilon}{4}(1 + \sqrt{1 + \frac{2}{\omega^2\upsilon}})$.

2.2 The score and GOF tests

When $\upsilon = 0$, the IGM distribution reduces to the $\mathtt{IG}(m, \omega)$ distribution. Desmond and Yang (2011) propose a score test of goodness of fit of IG distribution by testing $H_0 : \upsilon = 0$ vs $H_a : \upsilon > 0$, based on a complete sample $\{T_1, T_2, \ldots, T_n\}$. The test takes the form:

$$\mathtt{SC}_{\mathtt{IG}} = \left(\frac{\tilde\lambda\tilde\delta}{6n}\right)^{1/2} \sum_{i=1}^{n}\left(\tilde\lambda(\tilde\delta T_i - 1)^2 - T_i\right), \tag{2.9}$$

where, as in Desmond and Yang (2011), we adopt the parameterization in (2.3), and $\tilde\delta = \bar T^{-1}$ and $\tilde\lambda = (\tilde T - \bar T^{-1})^{-1}$ are the constrained MLEs of δ and λ under H_0, with $\bar T$ and $\tilde T$ denoting, respectively, the arithmetic and harmonic means of $\{T_1, T_2, \ldots, T_n\}$. The limiting null distribution of $\mathtt{SC}_{\mathtt{IG}}$ is $N(0, 1)$. Desmond and Yang (2011) argued that the test takes the same form when the alternative is the IGM given in (2.8), which is obtained from an arbitrary mixing distribution with mean m and variance υ satisfying a mild condition. Thus, the test is valid against various $\texttt{IGM}$ alternatives, including those given above.

Ducharme (2001) proposes a general purpose goodness of fit (GOF) test of inverse Gaussian distribution: $R_K = V_2 + \ldots, V_K$, where each V_K is asymptotically independent χ_1^2 under H_0 : T_i iid $\texttt{IG}(\delta, \lambda)$, and recommended the use of R_3. The terms V_2 and V_3 take the forms:

$$\begin{aligned} V_2 &= \tfrac{n\hat{\varphi}^4}{24+6\hat{\varphi}}\left(\bar{Z}_2 - \left(1 + \tfrac{3}{\hat{\varphi}} + \tfrac{3}{\hat{\varphi}^2}\right)\right)^2, \\ V_3 &= \tfrac{n\hat{\varphi}^6}{a(\hat{\varphi})}\left(\bar{Z}_3(4+\hat{\varphi}) - \bar{Z}_2\left(\tfrac{60}{\hat{\varphi}} + 30 + 4\hat{\varphi}\right) + b(\hat{\varphi})\right)^2, \end{aligned}$$

where $\bar{Z}_r = n^{-1}\sum_{i=1}^n (\bar{T}/T_i)^r$, $\hat{\varphi}$ is the constrained MLE of $\varphi = \delta\lambda$ under H_0, $a(\varphi) = 24(4+\varphi)(120+75\varphi+15\varphi^2+\varphi^3)$, and $b(\varphi) = 120\varphi^{-3}+195\varphi^{-2}+123\varphi^{-1}+32+3\varphi$. The limiting null distribution of R_3 is thus χ_2^2. For general principles on score tests, see Cox and Hinkley (1974); for general principles on GOF tests, see Rayner et al. (2009).

2.3 The score and GOF tests based on bootstrap critical values

The standard testing procedure approximates the finite sample null distribution of $\texttt{SC}_{\texttt{IG}}$ by $N(0,1)$ and the finite sample critical values by those of $N(0,1)$, say z_α, for $\alpha = 0.1, 0.05, 0.01$. However, the finite sample null distribution may be 'far' from $N(0,1)$, and thus its finite sample critical values may be 'far' from these of $N(0,1)$, leading to tests with large size distortions. The more skewed the $\texttt{IG}$ distribution, the larger the size distortion; also the heavier the censorship, the larger the size distortion as well. Hence, it is highly desirable to have methods that approximate better the finite sample critical values of the test statistic. We suggest the following simple procedure to obtain the bootstrap critical values:

(i) Compute $(\ddot{m}, \ddot{\omega})$: $\sqrt{n}$-consistent estimators of (m, ω) whether or not H_0 is true;

(ii) Draw a random sample $(T_1^b, T_2^b, \ldots, T_n^b)$ from $\texttt{IG}(\ddot{m}, \ddot{\omega})$;

(iii) Compute $\texttt{SC}_{\texttt{IG}}$ based on $(T_1^b, T_2^b, \ldots, T_n^b)$, and denote the resulting value by $\texttt{SC}_{\texttt{IG}}^b$;

(iv) Repeat (i)-(iii) B times, and compute the sample upper α-quantile, $q_\alpha(\ddot{m}, \ddot{\omega})$, of $\{\texttt{SC}_{\texttt{IG}}^b,\ b = 1, 2, \ldots, B\}$, which gives the bootstrap critical value for $\texttt{SC}_{\texttt{IG}}$ at H_0, or $\texttt{SC}_{\texttt{IG}}|_{H_0}$.

Yang (2015) points out that in bootstrapping the critical values of a score statistic, it is important to use parameter estimates that are consistent whether or not H_0 is true. This is because in real applications one does not know whether H_0 is true or not. In order to be able to mimic the parent distribution at the null, $\texttt{IG}(m, \omega)$, the distribution used to generate the bootstrap sample, $\texttt{IG}(\ddot{m}, \ddot{\omega})$, must converge to $\texttt{IG}(m, \omega)$ as n gets large, whether or not H_0 is true, and thus it is necessary that $(\ddot{m}, \ddot{\omega})$ be consistent for (m, ω) whether or not H_0 is true. With this requirement, it is clear that the choice of constrained estimates $(\tilde{m}, \tilde{\omega})$, under H_0, is not valid as they may not be consistent for (m, ω) when H_0 is false. Thus, unconstrained estimates $(\hat{m}, \hat{\omega})$, say, based on an $\texttt{IGM}$ distribution, given in (2.6)-(2.8), should be used. The formal argument is in Section 2.4 below.

Finite sample performance of Ducharme's GOF test R_3 is much worse than that of $\mathtt{SC_{IG}}$, with larger size distortions (see Table 2 of this paper) and lower size-adjusted power (see Tables IV and V of Desmond and Yang (2011)). Ducharme (2001) provides a table method for approximating the finite sample critical values of R_3. We show that an identical procedure as above for the $\mathtt{SC_{IG}}$ test can be followed to give bootstrap critical values of R_3 that essentially does an equivalent job to Ducharme's method (see Table 2 in Section 5). However, our method can easily be extended to the case of censored data, and the case of censored IG regression, to give asymptotically refined score tests, whereas Ducharme's GOF test does not seem to be extendible in a straightforward manner.

2.4 Validity of the bootstrap method

We now provide some formal arguments to show that with a proper choice of $(\ddot{m}, \ddot{\omega})$, $q_\alpha(\ddot{m}, \ddot{\omega})$ is able to provide a second-order approximation to $q_\alpha(m, \omega)$, the finite sample α-quantile of $\mathtt{SC_{IG}}|_{H_0}$, i.e., the error of approximation is of order $O(n^{-1})$. In contrast, the asymptotic critical value z_α gives only a first-order approximation to $q_\alpha(m, \omega)$, i.e., the error of approximation is of order $O(n^{-1/2})$. Let $\theta = (m, \omega)'$. Denote by $\mathcal{F}_n(\cdot, \theta)$ the finite sample null CDF of $\mathtt{SC_{IG}}$. Let $\boldsymbol{\kappa}_{j,n} \equiv \boldsymbol{\kappa}_{j,n}(\theta)$ be the jth cumulant of $\mathtt{SC_{IG}}|_{H_0}$. If $\boldsymbol{\kappa}_{4,n}$ exists, and $\boldsymbol{\kappa}_{j,n}$ can be expanded by a power series expansion in n^{-1}:

$$\boldsymbol{\kappa}_{j,n} = n^{-\frac{(j-2)}{2}}(k_{j,1} + n^{-1}k_{j,2} + n^{-2}k_{j,3} + \ldots),\ j = 1, 2, 3,$$

where $k_{1,1} = 0$ and $k_{2,1} = 1$, then $\mathcal{F}_n(\cdot, \theta)$ admits the following asymptotic expansion:

$$\mathcal{F}_n(s, \theta) = \Phi(s) + n^{-\frac{1}{2}}\phi(s)p(s, \theta) + O(n^{-1}), \tag{2.10}$$

where $p(s, \theta) = -k_{1,2} + \frac{1}{6}k_{3,1}(1 - s^2)$. See Hall (1992, p.46-48), or Yang (2015, Lemma A.7) for some more details. See also Hall and Horowitz (1996) for the application of the methods in the GMM context. From the expansion for $\boldsymbol{\kappa}_{j,n}$, we see that $k_{1,2} = \lim_{n\to\infty} \sqrt{n}\boldsymbol{\kappa}_{1,n}$, and $k_{3,1} = \lim_{n\to\infty} \sqrt{n}\boldsymbol{\kappa}_{3,n}$. Note that $k_{1,1} = 0$ and $k_{2,1} = 1$ correspond to the fact that the limiting mean and variance of $\mathtt{SC_{IG}}$ are 0 and 1.

Similar to the result that $\mathtt{SC_{IG}} \xrightarrow{D} N(0, 1)$, the bootstrap analogue $\mathtt{SC}^b_{\mathtt{IG}}$ of $\mathtt{SC_{IG}}$ is such that $\mathtt{SC}^b_{\mathtt{IG}} \xrightarrow{D^b} N(0, 1)$, where D^b denotes that the convergence is with respect to the bootstrap distribution, i.e., $\mathtt{IG}(\ddot{\mu}, \ddot{\omega})$, the estimated $\mathtt{IG}(\mu, \omega)$. Furthermore, the bootstrap CDF of $\mathtt{SC}^b_{\mathtt{IG}}$ must be of the form $\mathcal{F}_n(\cdot, \ddot{\theta})$, and similar to (2.10), admits the following expansion:

$$\mathcal{F}_n(s, \ddot{\theta}) = \Phi(s) + n^{-\frac{1}{2}}\phi(s)p(s, \ddot{\theta}) + O_p(n^{-1}). \tag{2.11}$$

Taking the difference between (2.10) and (2.11), we obtain

$$\mathcal{F}_n(s, \ddot{\theta}) - \mathcal{F}_n(s, \theta) = n^{-\frac{1}{2}}\phi(s)[p(s, \ddot{\theta}) - p(s, \theta)] + O_p(n^{-1}),$$

which is clearly of the order $O_p(n^{-1})$ if $\ddot{\theta}$ is a $\sqrt{n}$-consistent estimator of θ whether or not the null is true. This explains why the tests based on the proposed bootstrap critical values are able to provide a second-order approximation to the rejection probability, compared with the tests referring to the asymptotic critical values which give only a first order approximation to the rejection probability. The Monte Carlo results given in Section 5 show that using the constrained estimate $\tilde{\theta}$ leads to bootstrap critical values that are unstable, which can be significantly different from the 'true' ones, whereas using the unconstrained estimate $\hat{\theta}$ leads to very stable, close to truth, bootstrap critical values.

We note that for Ducharme's GOF test, the test statistic depends only on $\varphi = m\omega$, suggesting that the finite sample null CDF of R_3 depends only on φ. Hence, for generating bootstrap samples from the estimated $\texttt{IG}(m, \omega)$ distribution, it is only necessary to have an estimator of φ that is $\sqrt{n}$-consistent whether or not H_0 is true. Very interestingly, we see from the Monte Carlo results given in Section 5, that the bootstrap critical values of Ducharme's GOF test is quite insensitive to the choice of estimators of φ.

3 Score Tests Based on a Censored Sample

In this section, and the one which follows, we show how the score test considered in Section 2 can be easily extended to the case of a single right-censored sample, and the case of an inverse Gaussian regression model with complete or right-censored data. It is difficult, if possible at all, to derive the expected information-based score test, and the Hessian-based score test may have a complicated expression. We choose to present score tests based on outer-product-of-gradients (OPG), taking advantage of the fact that the observations are independent and hence the score functions can be written as sums of independent quantities. Again, we are concerned with the finite sample performance of the score tests. As noted above, to implement the bootstrap methods of Yang (2015), all that is needed is that the test statistic is an asymptotic pivot under the null, and that the estimators of the nuisance parameters (the parameters other than υ) are consistent whether or not the null is true.

Let $\mathbf{t} = \{t_1, t_2, \ldots, t_n\}$ denote the observed values and $\mathbf{r} = \{r_1, r_2, \ldots, r_n\}$ be the censoring indicators, i.e., $r_i = 1$ indicates that t_i is the observed failure time and $r_i = 0$ indicates it is a censored time. As for the case of a complete sample, derivation of the score test can simply be done based on the defective $\texttt{IGM}$ given in (2.4), as all the other proper IGM distributions lead to the same expression when letting $\upsilon \to 0$. The loglikelihood is, dropping the constants,

$$\ell(\theta|\mathbf{t}, \mathbf{r}) = n_r \log(\omega) - \frac{1}{2}\sum_{i=1}^{n} r_i \log(\upsilon t_i + 1) - \frac{1}{2}\sum_{i=1}^{n} \frac{r_i(mt_i - \omega)^2}{t_i(\upsilon t_i + 1)} + \sum_{i=1}^{n}(1 - r_i)\log \bar{G}(t_i|\theta), \quad (3.1)$$

where $\theta = (\omega, m, \upsilon)'$, $n_r = \sum_{i=1}^{n} r_i$, and $\bar{G}(t|\theta) = 1 - G(t|\theta)$; see Lawless (2003). The score functions are

$$
\begin{aligned}
S_\omega(\theta) &= \sum_{i=1}^n \left(\frac{r_i}{\omega} + \frac{r_i(mt_i-\omega)}{t_i(\upsilon t_i+1)} + (1-r_i)\frac{\bar{G}_\omega(t_i|\theta)}{\bar{G}(t_i|\theta)}\right) \equiv \sum_{i=1}^n s_{\omega,i}(\theta),\\
S_m(\theta) &= \sum_{i=1}^n \left(-\frac{r_i(mt_i-\omega)}{\upsilon t_i+1} + (1-r_i)\frac{\bar{G}_m(t_i|\theta)}{\bar{G}(t_i|\theta)}\right) \equiv \sum_{i=1}^n s_{m,i}(\theta),\\
S_\upsilon(\theta) &= \sum_{i=1}^n \left(-\frac{r_i t_i}{2(\upsilon t_i+1)} + \frac{r_i(mt_i-\omega)^2}{2(\upsilon t_i+1)^2} + (1-r_i)\frac{\bar{G}_\upsilon(t_i|\theta)}{\bar{G}(t_i|\theta)}\right) \equiv \sum_{i=1}^n s_{\upsilon,i}(\theta),
\end{aligned}
$$

where $\bar{G}_\varpi(t_i|\theta)$ denotes the partial derivatives of $\bar{G}(t_i|\theta)$ with respect to $\varpi = \omega, m$, and υ, respectively ; their exact expressions can easily be obtained and are given in the Appendix. Define an $n \times 3$ matrix $\mathcal{G}(\theta)$ with its ith row being $(s_{\omega,i}(\theta), s_{m,i}(\theta), s_{\upsilon,i}(\theta))$. Let $\tilde{\omega}$ and $\tilde{m}$ be the estimates of ω and m under $H_0 : \upsilon = 0$, and $\tilde{\theta} = (\tilde{\omega}, \tilde{m}, 0)'$. Let $\tau(\theta)$ be the square root of the bottom-right corner element of $[\mathcal{G}'(\theta)\mathcal{G}(\theta)]^{-1}$. Then, the score test for testing $H_0 : \upsilon = 0$, based on a censored sample takes the form:

$$
\mathtt{SC}^*_{\mathtt{IG}} = S_\upsilon(\tilde{\theta})\tau(\tilde{\theta}). \tag{3.2}
$$

The asymptotic null distribution of $\mathtt{SC}^*_{\mathtt{IG}}$ is $N(0,1)$, and the standard test is to refer to the standard normal critical values. These critical values give first-order approximations to the finite sample critical values of $\mathtt{SC}^*_{\mathtt{IG}}$ of which the exact finite sample distribution is unknown.

Bootstrap critical values. Censoring complicates the procedure for bootstrapping the score test statistic, as in practice one knows neither the censoring distribution nor whether or not H_0 is true. This is in a clear contrast to bootstrapping censored data for estimation or confidence interval construction where one faces the 'correct' model, and, if the censoring variable is independent of the failure time variable, a general and yet simple semiparametric procedure (with censoring distribution nonparametrically estimated) can be followed:

(1) Draw a random sample $(T_1^b, T_2^b, \ldots, T_n^b)$ from $\mathtt{IG}(\ddot{m}, \ddot{\omega})$;

(2) Sort the observed lifetimes as $(t_{(1)}, t_{(2)}, \ldots, t_{(n)})$, and denote the corresponding failure indicators by $(r_{(1)}, r_{(2)}, \ldots, r_{(n)})$;

(3) If $r_{(i)} = 0$, then the bootstrap censoring time $C_i^b = t_{(i)}$; otherwise C_i^b is a value randomly chosen from $(t_{(i)}, ..., t_{(n)})$;

(4) For $i = 1, \ldots, n$, set $t_i^b = \min(T_i^b, C_i^b)$, and $r_i^b = 1$ if $T_i^b < C_i^b$ otherwise 0.

See Davison and Hinkley (1997, Sec. 3.5) on this procedure. See also Efron (1981) for the very original idea on use of the bootstrap with censored data.

However, this procedure cannot be directly followed as steps (2) and (3) only mimic the true model (which may correspond to H_a). To bootstrap the finite sample critical value of the score test statistic it is necessary that the bootstrap procedure be able to mimic the null model whether it is true or not. To overcome this difficulty, we propose to *map* (or convert) the original data to a data set that is conformable with the null model. The details are as follows.

`Data Mapping Method:` *compute* $\hat{p}_i = G(t_i|\hat{m}, \hat{\omega}, \hat{v})$ *and* $\hat{t}_i = F^{-1}(\hat{p}_i|\hat{m}, \hat{\omega})$, $i = 1, 2, \ldots, n$, *and then replace the original data used in steps (2)-(3) above by the mapped data* $(\hat{t}_i, r_i)$, *where* $(\hat{m}, \hat{\omega}, \hat{v})$ *are the full estimates, and* F *and* G *are the* `IG` *and* `IGM` *CDFs given in Sec. 2.*

The above algorithm can be simplified if knowledge about the censoring distribution is available. For example, if it is known that censoring is fixed at one time point, say c_0, then the censoring time used to generate bootstrap samples is $\hat{c} = F^{-1}(\hat{p}|\ddot{m}, \ddot{\omega})$ where $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} r_i$; if it is known that the censoring is done uniformly around c_0, say $U(0.8c_0, 1.2c_0)$, then the bootstrap censoring times can be n random draws from $U(0.8\hat{c}, 1.2\hat{c})$. Monte Carlo results presented in Section 5 show these procedures perform well.

Clearly, the proposed bootstrap procedure makes the censoring proportions under the 'true' world and the 'null' world conformable to each other. This is important as a valid bootstrap procedure should be one that is able to mimic the 'null' world even if it may not represent the true one. Based on this method for generating a censored bootstrap sample, the detailed procedure for obtaining the bootstrap critical values for $\mathtt{SC}^*_{\mathtt{IG}}$ is as follows:

(i) Compute $(\ddot{m}, \ddot{\omega})$: $\sqrt{n}$-consistent estimators of (m, ω) whether or not H_0 is true;

(ii) Draw a censored bootstrap sample $\{(t_i^b, r_i^b), i = 1, \ldots, n\}$ based on $\mathtt{IG}(\ddot{m}, \ddot{\omega})$ according to the procedure described above;

(iii) Compute $\mathtt{SC}^*_{\mathtt{IG}}$ based on $\{(t_i^b, r_i^b), i = 1, \ldots, n\}$ and denote the resulting value by $\mathtt{SC}^{*b}_{\mathtt{IG}}$;

(iv) Repeat (i)-(iii) B times, and compute the sample upper α-quantile, $q^*_\alpha(\ddot{m}, \ddot{\omega})$, of $\{\mathtt{SC}^{*b}_{\mathtt{IG}},\ b = 1, 2, \ldots, B\}$, which gives the bootstrap critical value for $\mathtt{SC}^*_{\mathtt{IG}}|_{H_0}$.

It can be argued in a similar manner that the bootstrap critical value $q^*_\alpha(\ddot{m}, \ddot{\omega})$ is able to provide a second-order approximation to the finite sample critical value of $\mathtt{SC}^*_{\mathtt{IG}}|_{H_0}$.

From the above discussions, we see that it is indeed quite straightforward to extend the bootstrap score test to the case of a censored random sample. However, such an extension for Ducharme's GOF test is not straightforward. Ducharme (2001) describes a bootstrap test for Type I censored data, which is a special case of our random censoring scheme. It is not clear how to extend it to the general random censoring scheme. Further, the suggested parametric bootstrap method is questionable in the way the censoring times for the bootstrap samples are chosen, and in that the test is constructed by bootstrapping the score function and its variance. The score function itself is not an asymptotic pivot, and hence such a bootstrap procedure cannot achieve a second-order approximation. To construct a score pivot, it requires the explicit form of the survivor function under the alternative as seen above.

The issue left is how to obtain $\sqrt{n}$-consistent estimators of ω and m without knowing whether or not H_0 is true, to give a valid bootstrap test. If the class of the models considered are `IGM`, then general consistent estimators for ω and m can be obtained by maximizing the `IGM` likelihood. However, if the alternatives do not nest the null model as in the general purpose GOF test of Ducharme (2001), a different method is required to give generally consistent estimators of ω and m in order for the bootstrap method to be valid.

4 Score Tests for Inverse Gaussian Regression Model

4.1 The IG Regression Model

Moving from a constant drift to a random drift is one step accounting for the heterogeneity among individuals. Individuals may also differ in their initial health status (ω), and in their proclivity to drift to or away from the boundary, depending on their intrinsic health-related characteristics. Thus, as in the regular linear regression or generalized linear regression models, it is natural to assume that the mean (drift) of the underlying stochastic process is affected by some covariates. It is also natural to assume that the initial health status is affected by covariates. Consider a set of n randomly selected individuals with health-related characteristics (covariates) $X_1, \ldots, X_k$, initial values $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)'$, and drifts $\mathbf{m} = (m_1, \ldots, m_n)'$. As $\omega_i > 0$ and $m_i \geq 0$, plausible ways that link $\boldsymbol{\omega}$ and $\mathbf{m}$ to the covariates are:

$$\begin{aligned} \ln(\boldsymbol{\omega}) &= \gamma_0 1_n + \gamma_1 X_1 + \ldots + \gamma_k X_k = X\gamma, & (4.1) \\ \ln(\mathbf{m}) &= \beta_0 1_n + \beta_1 X_1 + \ldots + \beta_k X_k = X\beta, & (4.2) \end{aligned}$$

where 1_n is an n-vector of ones. Now, the test of $H_0 : \upsilon = 0$ corresponds to whether the IG regression model gives an adequate fit to the data or whether an alternative IGM regression model should be used. The link functions in (4.1) and (4.2) can be of other forms, and it is not necessary that they contain the same set of covariates.

4.2 Score tests based on complete data

Let $\theta = (\gamma', \beta', \upsilon)'$. The loglikelihood function is, omitting the constant terms,

$$\ell(\theta|\mathbf{t}, X) = \sum_{i=1}^{n} \log(\omega_i) - \frac{1}{2} \sum_{i=1}^{n} \log(\upsilon t_i + 1) - \frac{1}{2} \sum_{i=1}^{n} \frac{(m_i t_i - \omega_i)^2}{t_i(\upsilon t_i + 1)}, \tag{4.3}$$

where $\omega_i = \exp(x_i'\gamma)$, $m_i = \exp(x_i'\beta)$, and x_i' is the ith row of X. The score functions are

$$\begin{aligned} S_\gamma(\theta) &= \sum_{i=1}^{n} \left(1 + \frac{(m_i t_i - \omega_i)\omega_i}{t_i(\upsilon t_i + 1)}\right) x_i \equiv \sum_{i=1}^{n} s_{\gamma,i}(\theta), & (4.4) \\ S_\beta(\theta) &= \sum_{i=1}^{n} \left(-\frac{(m_i t_i - \omega_i)m_i}{\upsilon t_i + 1}\right) x_i \equiv \sum_{i=1}^{n} s_{\beta,i}(\theta), & (4.5) \\ S_\upsilon(\theta) &= \sum_{i=1}^{n} \frac{1}{2} \left(\left(\frac{m_i t_i - \omega_i}{\upsilon t_i + 1}\right)^2 - \frac{t_i}{\upsilon t_i + 1}\right) \equiv \sum_{i=1}^{n} s_{\upsilon,i}(\theta). & (4.6) \end{aligned}$$

Let $\tilde{\gamma}$ and $\tilde{\beta}$ be the MLEs of γ and β under the null, and $\tilde{\theta} = (\tilde{\gamma}', \tilde{\beta}, 0)'$. Let $\mathcal{G}(\theta)$ be an $n \times [2(k+1)+1]$ matrix with its ith row being $\{s_{\gamma,i}(\theta)', s_{\beta,i}(\theta)', s_{\upsilon,i}(\theta)\}$ and $\tau(\theta)$ be the square root of the bottom-right corner element of $[\mathcal{G}(\theta)'\mathcal{G}(\theta)]^{-1}$. The score test is

$$\mathtt{SC}_{\mathtt{IGR}} = S_\upsilon(\tilde{\theta})\tau(\tilde{\theta}). \tag{4.7}$$

The asymptotic null distribution of $\mathtt{SC}_{\mathtt{IGR}}$ is $N(0,1)$, and the test refers to the asymptotic critical values Z_α.

Bootstrap critical values. For IG regressions with complete data, the procedure for bootstrapping the critical values is a straightforward extension of that based on a complete sample. The details are as follows.

(i) Compute $(\ddot{\gamma}, \ddot{\beta})$: $\sqrt{n}$-consistent estimators of (γ, β) whether or not H_0 is true;

(ii) Draw t_i^b from $\mathtt{IG}(\ddot{m}_i, \ddot{\omega}_i)$, where $\ddot{m}_i = \exp(x_i'\ddot{\gamma})$ and $\ddot{\omega}_i = \exp(x_i'\ddot{\beta})$, independently for $i = 1, \ldots, n$;

(iii) Compute $\mathtt{SC}_{\mathtt{IGR}}$ based on $\{(t_i^b, x_i), i = 1, \ldots, n\}$, and denote the resulting value by $\mathtt{SC}^b_{\mathtt{IGR}}$;

(iv) Repeat (i)-(iii) B times, and compute the sample upper α-quantile, $q_\alpha(\ddot{\gamma}, \ddot{\beta})$, of $\{\mathtt{SC}^b_{\mathtt{IGR}},\ b = 1, 2, \ldots, B\}$, which gives the bootstrap critical value for $\mathtt{SC}_{\mathtt{IGR}}|_{H_0}$.

It can be argued in a similar manner that the bootstrap critical value $q_\alpha(\ddot{\gamma}, \ddot{\beta})$ is able to provide a second-order approximation to the finite sample critical value of $\mathtt{SC}_{\mathtt{IGR}}$.

4.3 Score tests based on right-censored data

With the censored regression data $\{(t_i, r_i, x_i), i = 1, \ldots, n\}$, the loglikelihood is, omitting the constants,

$$\ell(\theta|\mathbf{t}, \mathbf{r}, \mathbf{X}) = \sum_{i=1}^n r_i \left(\log(\omega_i) - \frac{1}{2}\log(\upsilon t_i + 1) - \frac{1}{2}\frac{(m_i t_i - \omega_i)^2}{t_i(\upsilon t_i + 1)} \right) + \sum_{i=1}^n (1 - r_i)\log \bar{G}_i(t_i|\theta), \tag{4.8}$$

where $\bar{G}_i(t_i|\theta) = \varphi[-a_i(t_i,\theta)] - c_i(\theta)\Phi[-b_i(t_i,\theta)]$, with $a_i(t_i,\theta) = \frac{m_i(\beta)t_i - \omega_i(\gamma)}{(\upsilon t_i^2 + t_i)^{1/2}}$, $b_i(t_i,\theta) = \frac{\omega_i(\gamma) + 2\omega_i(\gamma)\upsilon t_i + m_i(\beta)t_i}{(\upsilon t_i^2 + t_i)^{1/2}}$, and $c_i(\theta) = \exp[2\omega_i(\gamma)m_i(\beta) + 2\omega_i^2(\gamma)\upsilon]$. The score functions can simply be written as:

$$\begin{aligned}
S^*_\gamma(\theta) &= \sum_{i=1}^n r_i s_{\gamma,i}(\theta) + \sum_{i=1}^n (1 - r_i)\frac{\bar{G}_{i,\gamma}(t_i|\theta)}{\bar{G}_i(t_i|\theta)} \equiv \sum_{i=1}^n s^*_{\gamma,i}(\theta), \\
S^*_\beta(\theta) &= \sum_{i=1}^n r_i s_{\beta,i}(\theta) + \sum_{i=1}^n (1 - r_i)\frac{\bar{G}_{i,\beta}(t_i|\theta)}{\bar{G}_i(t_i|\theta)} \equiv \sum_{i=1}^n s^*_{\beta,i}(\theta), \\
S^*_\upsilon(\theta) &= \sum_{i=1}^n r_i s_{\upsilon,i}(\theta) + \sum_{i=1}^n (1 - r_i)\frac{\bar{G}_{i,\upsilon}(t_i|\theta)}{\bar{G}_i(t_i|\theta)} \equiv \sum_{i=1}^n s^*_{\upsilon,i}(\theta),
\end{aligned}$$

where $\bar{G}_{i,\varpi}(t_i|\theta) = \frac{\partial}{\partial\varpi}\bar{G}_i(t_i|\theta)$, for $\varpi = \gamma, \beta$ and υ; their expressions are given in the Appendix. Let $\tilde{\gamma}$ and $\tilde{\beta}$ be the MLEs of γ and β under the null, and $\tilde{\theta} = (\tilde{\gamma}', \tilde{\beta}, 0)'$. Let $\mathcal{G}^*(\theta)$ be an $n \times [2(k+1)+1]$ matrix with its ith row being $\{s^*_{\gamma,i}(\theta)', s^*_{\beta,i}(\theta)', s^*_{\upsilon,i}(\theta)\}$ and $\tau^*(\theta)$ be the square root of the bottom-right corner element of $[\mathcal{G}^*(\theta)'\mathcal{G}^*(\theta)]^{-1}$. The score test is

$$\mathtt{SC}^*_{\mathtt{IGR}} = S^*_\upsilon(\tilde{\theta})\tau^*(\tilde{\theta}), \tag{4.9}$$

of which the null distribution is asymptotically $N(0,1)$. However, the finite sample null behavior of $\mathtt{SC}^*_{\mathtt{IGR}}$ is unclear.

Bootstrap critical values. Generating bootstrap samples in a censored IG regression is trickier. The simplest way is to treat the censoring distribution to be independent of the covariates, besides being independent of the failure times. In this case, the bootstrap censoring times can be generated in a similar manner as that described in the `data mapping method` given in Section 3 for one sample problems. Now to map the original data (t_i, r_i) into the null world in a regression framework, compute $\hat{p}_i = G(t_i|\hat{m}_i, \hat{\omega}_i, \hat{\upsilon})$, and $\hat{t}_i = F^{-1}(\hat{p}_i|\hat{m}_i, \hat{\omega}_i)$, $i = 1, 2, \ldots, n$. The mapped data are thus $\{(\hat{t}_i, r_i), i = 1, \ldots, n\}$. The details for generating bootstrap data are:

(1) Draw T_i^b from $\mathtt{IG}(\ddot{m}_i, \ddot{\omega}_i)$, independently for $i = 1, \ldots, n$;

(2) Sort the mapped lifetimes in ascending order, and denote the results by $(\hat{t}_{(1)}, \hat{t}_{(2)}, \ldots, \hat{t}_{(n)})$ and the corresponding failure indicators by $(r_{(1)}, r_{(2)}, \ldots, r_{(n)})$;

(3) If $r_{(i)} = 0$, then the bootstrap censoring time $C_i^b = \tilde{t}_{(i)}$; otherwise C_i^b is a value randomly chosen from $(\hat{t}_{(i)}, \ldots, \hat{t}_{(n)})$;

(4) For $i = 1, \ldots, n$, set $t_i^b = \min(T_i^b, C_i^b)$, and $r_i^b = 1(T_i^b < C_i^b)$.

Now, based on this method for generating the censored bootstrap data, the procedure for bootstrapping the critical values of $\mathtt{SC}^*_{\mathtt{IGR}}|_{H_0}$ is as follows:

(i) Compute $(\ddot{\gamma}, \ddot{\beta})$: $\sqrt{n}$-consistent estimators of (γ, β) whether or not H_0 is true;

(ii) Obtain the bootstrap data $\{(t_i^b, r_i), 1, \ldots, n\}$ based on the procedure given above;

(iii) Compute $\mathtt{SC}^*_{\mathtt{IGR}}$ based on $\{(t_i^b, r_i, x_i), 1, \ldots, n\}$, and denote the resulting value by $\mathtt{SC}^{*,b}_{\mathtt{IGR}}$;

(iv) Repeat (i)-(iii) B times, and compute the sample upper α-quantile, $q^*_\alpha(\ddot{\gamma}, \ddot{\beta})$, of $\{\mathtt{SC}^{*,b}_{\mathtt{IGR}}, b = 1, 2, \ldots, B\}$, giving the bootstrap critical value for $\mathtt{SC}_{\mathtt{IGR}}|_{H_0}$.

It can be shown that if $\ddot{\gamma}$ and $\ddot{\beta}$ are consistent estimators of γ and β, $q^*_\alpha(\ddot{\gamma}, \ddot{\beta})$ offers a second-order approximation to $q^*_\alpha(\gamma, \beta)$, the unknown finite sample critical value of $\mathtt{SC}^*_{\mathtt{IGR}}|_{H_0}$. If the censoring distribution depends the covariates, the proposed data mapping method can still be applied as the null and the alternative correspond to the same set of covariates.

5 A Monte Carlo comparison

Extensive Monte Carlo experiments are run to investigate the finite sample performance of the tests when referring to the asymptotic critical values (ACRs), or the bootstrap critical values (BCRs) based on the restricted MLEs, or the BCRs based on the unrestricted estimates. As remarked in Yang (2015), in investigating the finite sample behavior of a bootstrap test it is important to see whether the bootstrap critical values are stable with respect to the alternatives, in our context, the value of υ. Thus, the stability of the two types of bootstrap critical values with respect to the change in the value of υ is also assessed. The methods for generating the IG and IGM random variates can be found in, e.g., Desmond and Yang (2011).

The case of a complete sample. Consider first the one-sample test with complete data. For simplicity, we use the defective IGM given in (2.4) to estimate jointly the parameters (m, ω, υ), and use the resulting MLEs $\hat{m}$ and $\hat{\omega}$ as the bootstrap parameters to generate bootstrap samples from the IG$(\hat{m}, \hat{\omega})$. We choose the parameter values so that the (2.4) is essentially a proper pdf so that $(\hat{m}, \hat{\omega})$ is $\sqrt{n}$-consistent for (m, ω) whether or not $\upsilon = 0$. The results are summarized in Tables 1a, 1b and 2a and 2b. Each set of results (corresponding to one parameter configuration) is based on $M = 5000$ Monte Carlo samples and $B = 999$ bootstrap samples from each Monte Carlo sample.

First from Table 1a we see that exactly as the theory predicts, the BCR for the score test based on the restricted MLE (rMLE) varies significantly with the alternative value υ. In contrast, the BCR for the score test based on the unconstrained MLE (uMLE) is very stable with respect to the change of the υ value. Furthermore, BCRs based on uMLE are very close to the 'true' finite sample critical values (FCRs), obtained from the Monte Carlo simulation based on 50,000 Monte Carlo runs. This confirms our theory that the bootstrap critical value is a second-order approximation to the finite sample critical values of the test statistic. The consequence of using rMLE, as seen clearly from the results in Table 2a, is that the power of the test is unreliable – it tends to give higher power due to the fact that the test is based on a smaller critical value (than the true one). The score test referring to ACRs can be severely undersized for the 90% and 95% tests, but oversized for the 1% test. In contrast, the test referring to BCRs have sizes very close to their nominal levels.

The results in Tables 1b and 2b show a similar pattern for Ducharme's GOF test as for the score test. Some noticeable differences are as follows. The GOF test referring to ACRs has much larger and much more persistent size distortions, showing the stronger need of using the bootstrap critical values. When the size of the test is adjusted (by referring to the appropriate BCRs), the GOF test is seen to have a significantly lower power than the score test. Again, the BCRs based on uMLE are more stable than those based on rMLE, and are very close to the corresponding FCRs obtained from Monte Carlo simulation.

Interestingly, we note that while the BCRs under uMLEs are very close to the corresponding FCRs for any sample size, which are correct from our theoretical point of view, the convergence of FCRs to the corresponding ACRs is very slow, in particular for the GOF test. In other words, the convergence at the null of the score test to $N(0,1)$ is very slow, and that of GOF test to χ_2^2 is even slower. These observations further reinforce the need to use BCRs in performing tests. The advantage of the bootstrap method is that it does not require the knowledge of the limiting distribution of the test statistic, as long as it is asymptotically pivotal.

The case of a censored sample. We consider various censoring schemes such as Type I, Type II, random censoring with known censoring distribution, and random censoring with unknown censoring distribution where the proposed `data mapping method` is applied. For each Monte Carlo samples, two sets of observations $\{T_1^*, \ldots, T_n^*\}$ and $\{C_1 \ldots, C_n\}$ are generated, with T_i^* from the 'true' distribution (IG or IGM), and C_i from any proper distribution referred to as the censoring distribution. Tables 3a and 3b summarize the bootstrap critical values and the rejection rates under the fixed-time censoring (Type I) scheme with c_0 being the 90th percentile of the true distribution, where the bootstrap censoring time is the estimate $\hat{c} = F^{-1}(\hat{p}|\ddot{m}, \ddot{\omega})$ with $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} r_i$ as described in Section 3. Tables 4a and 4b replicate the results in Tables 3a and 3b with c_0 changed to the 80th percentile of the true distribution. Tables 5a, 5b and 5c contain the bootstrap critical values and rejection rates where the censoring times are random draws from $U(0.8c_0, 1.2c_0)$ with c_0 being the 80th percentile of the true distribution. Table 5b corresponds to the case of a known censoring distribution, and in this case the bootstrap censoring times are random draws from $U(.8\hat{c}, 1.2\hat{c})$, whereas Table 5c corresponds to the case of an unknown censoring distribution, and the bootstrap censoring times are generated from the `data mapping method` described in Section 3. As computation under censored data is more demanding, we reduce the number of Monte Carlo samples to $M = 2000$ and the number of bootstrap samples to $B = 699$.

The results reveal a similar general pattern as in the case of complete data: (i) the score test referring to the asymptotic critical values can have large size distortion, (ii) bootstrap critical values are very close to the finite sample critical values (obtained from the Monte Carlo simulation), and the test referring to the bootstrap critical values has sizes very close to the nominal levels. Furthermore, comparing the results in Table 4b with these in Table 3b we see the heavier the censorship, the larger the size distortion of the asymptotic test. Again, the results show that the convergence of the score test under the null to $N(0,1)$ is very slow. Finally and very interestingly, the results in Table 5c show that the proposed `data mapping method` for generating bootstrap censoring times performs very well in terms of the size of the test, although it has slightly lower power than the case where the censoring distribution is treated as known (Table 5b). Results under Type II censoring are omitted due to space limitations.

IG regression with complete data. Consider the following models:

$$\ln(m_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \;\text{ and }\; \ln(\omega_i) = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i},\; i = 1, 2, \ldots, n,$$

where the values $\{x_{1i}\}$ and $\{x_{2i}\}$ of the two covariates X_1 and X_2 are generated independently from $N(0, .5^2)$ and are fixed in the subsequent Monte Carlo runs. The values of the γ-parameters are chosen to be $(.5, .25, .25)$ and so are the values of the β-parameters. In choosing the parameter values for an IG regression model, it may be helpful to recall that the mean and variance for an $\texttt{IG}(\mu, \omega)$ distribution are, respectively, ω/μ and ω/μ^3. This choice of parameter values thus gives $\delta_i = \mu_i/\omega_i = 1$ as in the one sample case, and $\lambda_i = \omega_i^2$ with a similar magnitude as in the one sample case. Monte Carlo results under other parameter values (unreported to conserve space) show the parameter values may affect the finite sample performance of the test. We use again $M = 2000$ and $B = 699$ as numerical maximization needs to be done for estimating any of the model parameters.

Table 6a contains Monte Carlo results corresponding to the finite sample distribution of the score test statistic $\text{SC}_{\texttt{IGR}}$ defined in (4.7). The results show that $\text{SC}_{\texttt{IGR}}$ can be very much negatively skewed, and that the convergence of $\text{SC}_{\texttt{IGR}}$ to $N(0, 1)$ is extremely slow and is much worse than the case of a complete sample. Interestingly, the positive part of $\text{SC}_{\texttt{IGR}}$ is much closer to $N(0, 1)$ than its negative part. This is a desirable feature as only right-tailed tests are concerned. However, there must be scenarios under which $\text{SC}_{\texttt{IGR}}$ becomes positively skewed. More research is desirable but is beyond the scope of this paper.

Table 6b contains the bootstrap critical values based on the restricted and the unrestricted parameter estimates. From the results we see that the bootstrap critical values are in general very close to the finite sample critical values. It is very interesting to note that, similar to the case of a single censored sample, the bootstrap critical values in the inverse Gaussian regression models with complete data are quite robust to the choice of the parameter estimates, although in general the bootstrap critical values based on the unrestricted estimates are more stable.

Table 6c contains the rejection probabilities for the IG regression model with complete data. From the results, we see that the size and power of the test based on the asymptotic critical values can be quite unreliable. Both size and power can be too small relative to the 'true' size and power, even when the sample size is very large (e.g., $n = 2000$, as the results of Table 6a suggest). In contrast, the test based on the bootstrap critical values are much more reliable.

IG regression with censored data. We consider the same models as in the case of complete data. Censoring makes the estimation of IGM regression model more difficult when n is not so large. Considering the fact that the bootstrap critical values are quite robust against the choice of parameter estimates, we focus on comparing the results based on ACR and BCR under rMLE. We consider various censoring schemes such as Type I, Type II and random censoring. The Type I censoring time is fixed at $c_0 =$ 90th percentile of $\texttt{IGM}(\bar{m}, \bar{\omega}, \upsilon)$, where $\bar{m}$ and $\bar{\omega}$ are the averages of m and ω, respectively; the bootstrap censoring time $\hat{c} = F^{-1}(\hat{p}|\bar{\tilde{m}}, \bar{\tilde{\omega}})$, where $\hat{p}$ is defined above, and $\bar{\tilde{m}}$ and $\bar{\tilde{\omega}}$ are the averages of the restricted estimates $\tilde{m}$ and $\tilde{\omega}$. In the case of random censoring, as the `data mapping` method described in Section 3 requires the estimation of the IGM regression model, we focus on the case of a known censoring distribution: the censoring times are random draws from $U(0.8c_0, 1.2c_0)$, and bootstrap censoring times are random draws from $U(0.8\hat{c}, 1.2\hat{c})$. For Type II censoring, the censoring time is the $(.9n)$th order statistic, and the bootstrap censoring time is the $(.9n)$th order statistic of the sample generated from $\texttt{IG}(\tilde{m}_i, \tilde{\omega}_i), i = 1, 2, \ldots, n$.

Tables 7a, 7b, and 7c contain partial results under Type II censoring, Tables 8a, 8b and 8c contain partial results under random censoring, and for the Type I censoring case, only the finite sample critical values are given in Table 9, as Type I censoring can be considered as a special case of random censoring. From the results, it is interesting to see that the finite sample null behaviour of the score test for `IG` regression under censoring is 'nicer' compared with the case of complete data – it is more symmetrically distributed and it converges to $N(0, 1)$ faster, in particular under Type I and random censorship. The bootstrap critical values based on the restricted parameter estimates are quite stable with respect to alternative values, and the tests referring to `BCRr` have much more reliable size and power.

6 Real Data Illustrations

In this section, we provide some real data illustrations on the practical implementations of the score and GOF tests referring to the asymptotic as well as the bootstrap critical values.

(1) **The repair time data** (Whitmore, 1986). The data consist of 46 repair times (in hours) for an airborne communication transceiver. The restricted and unrestricted estimates of (δ, λ) are both $(0.2773, 1.6589)$ as the unrestricted estimate $\hat{\upsilon}$ of υ is 0. The score test statistic $\mathrm{SC}_{\texttt{IG}} = -0.2899$, and the GOF test $R_3 = 0.0093$, both suggesting that the IG fits the data well. The table below gives the bootstrap critical values (`BCR`), showing that they may differ dramatically from the corresponding asymptotic critical values (`ACR`).

Bootstrap Critical Values: Repair Time Data

Nominal	Score Test		GOF Test	
size	ACR	BCR	ACR	BCR
10%	1.2816	0.4586	4.6052	2.8401
5%	1.6449	0.8468	5.9915	4.2406
1%	2.3263	2.0395	9.2103	9.5741

(2) **The Reaction Time Data** (Whitmore, 1986). Based on a complete sample of 180 reaction times with origin shifted by 337.5, the restricted estimates are $(\tilde{\delta}, \tilde{\lambda}) = (0.0026457, 1019.5)$ and the unrestricted estimates are $(\hat{\delta}, \hat{\lambda}, \hat{\upsilon}) = (0.0030534, 1744.8, 0.0017134)$. The score test with $\mathrm{SC}_{\texttt{IG}} = 3.6678$ and the GOF test with $R_3 = 14.0502$ both reject the null hypothesis of IG distribution, based on either the `ACR`s or the `BCR`s, shown in the table below.

Bootstrap Critical Values: Reaction Time Data

Nominal	Score Test			GOF Test		
size	ACR	BCRr	BCRu	ACR	BCRr	BCRu
10%	1.2816	1.3590	1.4692	4.6052	3.5703	3.8021
5%	1.6449	1.9740	2.0179	5.9915	5.4046	5.5016
1%	2.3263	3.2080	2.8182	9.2103	11.469	11.919

Even with a complete sample of size 180, we see that the bootstrap critical values differ significantly from the asymptotic ones. Use of restricted or unrestricted estimates as the bootstrap parameters also makes a noticeable difference on the `BCRs` for the score test, but not much for the GOF test. This is consistent with the Monte Carlo results presented earlier.

(3) **The Failure Age Data** (Whitmore, 1983). This is a censored data set with 20 observations (three censored) on the failure age of aluminum-reduction cells and one regressor, the time since the cell was first installed. For details, see Whitmore (1983). As in Whitmore (1983), we first analyze the failure ages treated as one sample, and then analyze the whole data using our proposed regression model. First, for the single-sample analysis, both restricted and unrestricted estimates of δ and λ are 77.2058 and 0.0479 as the unrestricted estimate of υ is 0. The score test statistic has a value of -0.7590, leading to non-rejection of the IG hypothesis. The `BCRs` are $(1.9611, 2.4404, 3.0520)$, which are much larger than the corresponding `ACRs`.

For the regression analysis we use Models (4.1) and (4.2) with the sole regressor and an intercept appearing in both terms. Both restricted and unrestricted estimates of $(\gamma_0, \gamma_1, \beta_0, \beta_1)$ are $(4.374, .0012388, -3.4142, .0086256)$ as the estimate of υ is 0. The score test statistic has a value of -2.1467, leading to a non-rejection of IG hypothesis. The `BCRs` are $(2.7596, 3.2072, 3.7156)$, significantly larger than the corresponding ACRs, showing the importance of using the bootstrap critical values for inference.

(4) **The Motorettes Failure Data** (Whitmore, 1983). This is a Type I censored regression data with 10 observations corresponding to each of the four levels of the sole regressor, which is temperature in degrees Celsius. The first 10 observations are all censored. We follow Whitmore (1983) to use the remaining 30 observations to perform a regression analysis. We use the same model as that in (3) above, which is different from the model used in Whitmore (1983). The restricted and unrestricted estimates of $(\gamma_0, \gamma_1, \beta_0, \beta_1)$ are $(4.5342, -.0015461, -11.471, .042144)$ and $(4.1385, .0038305, -12.489, .049489)$. The estimate of υ is .015343. The score statistic has a value of -2.0875, leading to non-rejection of the IG assumption. `BCRr` $= (2.14, 2.6578, 3.5585)$ and `BCRu` $= (3.1731, 3.8742, 5.2737)$. Unlike the case in (3), this is a heavily censored data set, causing the unrestricted estimation and the `BCRu` values to be sensitive to the initial values, because the true υ is at or close to the boundary value 0. In this case, inference based on `BCRr` is recommended, in line with the discussions given at the end of Sec. 5.

7 Conclusion and Discussion

Score tests for inverse Gaussian models are considered. Bootstrap methods are proposed to give a second-order approximation to the finite sample critical values of the score test statistic. Implementation of the bootstrap method in the complete data case is quite straightforward, but not in the censored data case. We propose a data mapping method that maps the real data into the null model to make sure the resulting censoring observations mimic those under the real world. This method is valid as long as the censoring distribution is independent of the lifetime distribution. Extensive Monte Carlo experiments demonstrate excellent performance of the proposed bootstrap score tests for inverse Gaussian distribution or inverse Gaussian regression. In the case of the former, it is advisable to use the unrestricted parameter estimates to generate bootstrap IG samples, whereas in the case of the latter, the bootstrap critical values are seen to be quite robust against the choice of parameter estimates and hence, considering its simplicity, the restricted estimates are recommended for use in generating bootstrap samples.

The score test is known to be a member of the usual 'Holy Trinity' of asymptotic likelihood-based tests. A reviewer has drawn our attention to a very interesting addition to this 'Holy Trinity', the gradient test (Terrell, 2002; Lemonte, 2012), which uses only the point estimate and the concentrated score of the parameter to be tested, based on a remarkably simple idea. While this test clearly deserves further attention as pointed out by Rao (2005), a detailed study is beyond the scope of the paper, and is left for future research. Instead, we have conducted some Monte Carlo experiments to study the size properties of all these four tests. To ensure comparability with the score and Wald tests, we have used the signed square-root version of the gradient test and the signed square-root likelihood ratio test. All of them are referred to the upper critical values of the standard normal distribution. The results (not reported for brevity) show that all these four tests may have poor finite sample properties, in particular the gradient test with censored data. The likelihood ratio test seems to have the best finite sample size property, but its size distortion can still be significant. This shows the need for asymptotic refinements for all these four tests. However, the calculation of the gradient test, as of the likelihood ratio and Wald tests, requires the estimation of the full model in every bootstrap sample, making the bootstrap approximation to the finite sample critical values extremely intensive computationally.

Another type of test which can be used in our situation is the Kolmogorov-Smirnov (KS) test (Edgeman,1990; Pavur et al., 1992). The KS test is an omnibus test often lacking power for specific directions. Our test is designed for specific alternatives important in the use of IG in applications in reliability and survival analysis, where unobserved heterogeneity in the drift parameter is important. The situation is analogous to frailty effects in more conventional survival models. Furthermore, the proposed bootstrap method requires consistent estimators (whether or not H_0 is true) to be used as bootstrap parameters. A clearly specified alternative leads to general consistent estimators.

Acknowledgments

The authors are grateful to an Associate Editor and two anonymous referees for their helpful comments. Anthony F. Desmond's research is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. Zhenlin Yang gratefully acknowledges the research support from Singapore Management University. He thanks the Dept. of Math. and Stat., University of Guelph, for the support and hospitality during his visit in Oct. 2014.

References

[1] Aalen, O. O. (1994). Effects of frailty in survival analysis. *Statistical Methods in Medical Research* 3, 227-243

[2] Aalen, O. O., Borgan, O. and Gjessing, H. K. (2008). *Survival and event history analysis: A process point of view.* New York: Springer.

[3] Banerjee, A. K. and Bhattacharyya, G. K. (1976). A purchase incidence model with inverse Gaussian interpurchase times. *Journal of the American Statistical Association*, **71**, 823-829.

[4] Chhikara, R. S. and Folks, J. L. (1977). The inverse Gaussian distribution as a lifetime model. *Technometrics*, **19**, 461-468.

[5] Chhikara, R. S. and Folks, J. L. (1978). The inverse Gaussian distribution and its statistical application - a review. *J. R. Statist. Soc.* B, **40**, 263-289.

[6] Chhikara, R. S. and Folks, J. L. (1989). *The inverse Gaussian distribution: theory, methodology and applications.* Marcel Dekker, New York and Basel.

[7] Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

[8] Davison A. C. and Hinkley D. V. (1997). *Bootstrap Methods and Their Applications.* Cambridge University Press: London.

[9] Desmond, A. F. and Chapman, G. R. (1993). Modelling task completion data with inverse Gaussian mixtures. *Applied Statistics*, **42**, 603-613.

[10] Desmond, A. F. and Yang, Z. L. (2011). Score tests for inverse Gaussian mixtures. *Applied Stochastic Models in Business and Industry* 27, 633-648.

[11] Ducharme, G. R. (2001). Goodness-of-fit tests for the inverse Gaussian and related distributions. *Test*, **10**, 271-290.

[12] Edgeman, R. L. (1989). Inverse Gaussian control charts. *Australian Journal of Statistics*, **31**, 78-84.

[13] Edgeman, R. L. (1990). Assessing the inverse Gaussian distribution assumption. *IEEE Transactions on Reliability.* **R39**, 352-355.

[14] Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association* **76**, 312-319.

[15] Feller, W. (1966). *An Introduction to Probability theory and its Applications, Vol. II.* Wiley. New York.

[16] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distribution.* Wiley: New York.

[17] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer, New York.

[18] Hall, P. and Horowitz, J. L. (1996). Bootstrap critical values for tests based on generalized-methods of moments estimators. *Econometrica* **64**, 891-916.

[19] Hawkins, D. M. and Olwell, D. H. (1997). Inverse Gaussian cumulative sum control charts for location and shape. *The Statistician*, **46**, 323-335.

[20] Lancaster, T. (1972). A stochastic model for the duration of a strike. *J. R. Statist. Soc.* A, **135**, 257-271.

[21] Lawless, J. F. (2003). *Statistical models and methods for lifetime data*, 2nd ed. John Wiley & Sons, Inc., Hoboken, New Jersey.

[22] Lee, M.-L. T. and Whitmore, G. A. (2006). Threshold regression for survival analysis: Modeling event times by a stochastic process reaching a boundary. *Statistical Science* 4, 501-513.

[23] Lee, M.-L. T. and Whitmore, G. A. (2010). Proportional hazards and threshold regression: Their theoretical and practical connections. *Lifetime Data Analysis* 16, 196-214.

[24] Lemonte, A. J. and Ferrari, S. L. P. (2012). The local power of the gradient test. *Annals of the Institute of Statistical Mathematics* 64, 373381.

[25] Letac, G., Seshadri, V. and Whitmore, G. A. (1985). An exact chi-squared decomposition theorem for inverse Gaussian variates. *J. R. Statist. Soc.* B, **47**, 476-481.

[26] Pavur, R. J., Edgeman R. L. and Scott, R. C. (1992). Quadratic statistics for the goodness-of-fit test of the inverse Gaussian distribution. *IEEE Transactions on Reliability*, **41**, 118-123.

[27] Rao, C. R. (2005). Score test: historical review and recent developments. In N. Balakrishnan, N. Kannan, H. N. Nagaraja (Eds.), *Advances in ranking and selection, multiple comparisons, and reliability.* Boston: Birkhauser.

[28] Rayner, J. C. W., Thas, O. and Best, D. J. (2009). *Smooth Tests of Goodness of Fit Using R, 2nd ed.* John Wiley & Sons, Singapore.

[29] Seshadri, V. (1993). *The inverse Gaussian distribution: A Case Study in Exponential Families.* Oxford University Press, Oxford.

[30] Seshadri, V. (1998). *The inverse Gaussian distribution.* Springer, New York.

[31] Terrell, G. R. (2002). The gradient statistic. *Computing Science and Statistics* 34, 206-215.

[32] Whitmore, G. A. (1979). An inverse Gaussian model for labour turnover. *J. R. Statist. Soc.* A, **142**, 468-478.

[33] Whitmore, G. A. (1983). A regression method for censored inverse-Gaussian data. *Canadian Journal of Statistics* 11, 305-315.

[34] Whitmore, G. A. (1986). Normal-gamma mixtures of inverse Gaussian distribution. *Scandinavian Journal of Statistics* 13, 211-220.

[35] Yang, Z. L. (2015). LM tests of spatial dependence based on bootstrap critical values. *Journal of Econometrics* 185, 33-59.

Appendix: Some Detailed Expressions

Some expressions in Sections 3-5 are quite straightforward to derive, but can be lengthy. We give them in this appendix to facilitate the practical applications of the proposed tests.

Expressions for Section 3. Write $G(t|\theta) = \Phi(a(t,\theta)) + c(\theta)\Phi(-b(t,\theta))$, where $a(t,\theta) = (mt-\omega)(\upsilon t^2+t)^{-1/2}$, $b(t,\theta) = (mt+\omega+2\omega\upsilon t)(\upsilon t^2+t)^{-1/2}$, and $c(\theta) = \exp(2\omega m + 2\omega^2\upsilon)$. We have, $\bar{G}(t|\theta) = \Phi(-a(t,\theta)) - c(\theta)\Phi(-b(t,\theta))$, and hence

$$\begin{aligned}
\bar{G}_\omega(t|\theta) &= -\phi(-a(t,\theta))a_\omega(t,\theta) - c_\omega(\theta)\Phi(-b(t,\theta)) + c(\theta)\phi(-b(t,\theta))b_\omega(t,\theta),\\
\bar{G}_m(t|\theta) &= -\phi(-a(t,\theta))a_m(t,\theta) - c_m(\theta)\Phi(-b(t,\theta)) + c(\theta)\phi(-b(t,\theta))b_m(t,\theta),\\
\bar{G}_\upsilon(t|\theta) &= -\phi(-a(t,\theta))a_\upsilon(t,\theta) - c_\upsilon(\theta)\Phi(-b(t,\theta)) + c(\theta)\phi(-b(t,\theta))b_\upsilon(t,\theta),
\end{aligned}$$

where

$a_\omega(t,\theta) = -(\upsilon t^2+t)^{-1/2}$,
$a_m(t,\theta) = t(\upsilon t^2+t)^{-1/2}$,
$a_\upsilon(t,\theta) = -\frac{1}{2}(mt-\omega)t^2(\upsilon t^2+t)^{-3/2}$,
$b_\omega(t,\theta) = (1+2\upsilon t)(\upsilon t^2+t)^{-1/2}$,
$b_m(t,\theta) = a_m(t,\theta),\ b_\upsilon(t,\theta) = -\frac{1}{2}(mt+\omega+2\omega\upsilon t)t^2(\upsilon t^2+t)^{-3/2} + 2\omega t(\upsilon t^2+t)^{-1/2}$.

Expressions for Section 4. Recall $\bar{G}_i(t_i|\theta) = \Phi[-a_i(t_i,\theta)] - c_i(\theta)\Phi[-b_i(t_i,\theta)]$, where $a_i(t_i,\theta) = \frac{m_i(\beta)t_i-\omega_i(\gamma)}{(\upsilon t_i^2+t_i)^{1/2}}$, $b_i(t_i,\theta) = \frac{\omega_i(\gamma)+2\omega_i(\gamma)\upsilon t_i+m_i(\beta)t_i}{(\upsilon t_i^2+t_i)^{1/2}}$, and $c_i(\theta) = \exp[2\omega_i(\gamma)m_i(\beta)+2\omega_i^2(\gamma)\upsilon]$. We obtain the partial derivatives of $\bar{G}_i(t_i|\theta)$ as,

$$\begin{aligned}
\bar{G}_{i,\gamma}(t_i|\theta) &= -\phi(-a_i(t_i,\theta))a_{i\gamma}(t_i,\theta) - c_{i\gamma}(\theta)\Phi(-b_i(t_i,\theta)) + c_i(\theta)\phi(-b_i(t_i,\theta))b_{i\gamma}(t,\theta),\\
\bar{G}_{i,\beta}(t_i|\theta) &= -\phi(-a_i(t_i,\theta))a_{i\beta}(t_i,\theta) - c_{i\beta}(\theta)\Phi(-b_i(t_i,\theta)) + c_i(\theta)\phi(-b_i(t_i,\theta))b_{i\beta}(t_i,\theta),\\
\bar{G}_{i,\upsilon}(t_i|\theta) &= -\phi(-a_i(t_i,\theta))a_{i\upsilon}(t_i,\theta) - c_{i\upsilon}(\theta)\Phi(-b_i(t_i,\theta)) + c_i(\theta)\phi(-b_i(t_i,\theta))b_{i\upsilon}(t_i,\theta),
\end{aligned}$$

where

$a_{i\gamma}(t_i,\theta) = -\frac{\omega_i(\gamma)}{(\upsilon t_i^2+t_i)^{1/2}}x_i$,
$a_{i\beta}(t_i,\theta) = \frac{m_i(\beta)t_i}{(\upsilon t_i^2+t)^{1/2}}x_i$,
$a_{i\upsilon}(t_i,\theta) = -\frac{[m_i(\beta)t_i-\omega_i(\gamma)]t_i^2}{2(\upsilon t_i^2+t_i)^{3/2}}$,
$b_{i\gamma}(t_i,\theta) = \frac{\omega_i(\gamma)(1+2\upsilon t_i)}{(\upsilon t_i^2+t_i)^{1/2}}x_i$,
$b_{i\beta}(t_i,\theta) = a_{i\beta}(t_i,\theta)$,
$b_{i\upsilon}(t_i,\theta) = \frac{2\omega_i(\gamma)t_i}{(\upsilon t_i^2+t_i)^{1/2}} - \frac{[m_i(\beta)t_i+\omega_i(\gamma)+2\omega_i(\gamma)\upsilon t_i]t_i^2}{2(\upsilon t_i^2+t_i)^{3/2}}$,
$c_{i\gamma}(\theta) = c_i(\theta)[2\omega_i(\gamma)m_i(\beta) + 4\omega_i^2(\gamma)\upsilon]x_i$,
$c_{i\beta}(\theta) = 2c_i(\theta)\omega_i(\gamma)m_i(\beta)x_i$, and
$c_{i\upsilon}(\theta) = 2c_i(\theta)\omega_i^2(\gamma)$.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 1a. Bootstrap Critical Values of Score Test for IG Distribution: $m = \omega = 2$

Uncensored Data

	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
10% Score Test: ACR = 1.2816										
.0	0.883	0.910	1.057	1.069	1.147	1.152	1.209	1.211	1.255	1.256
.1	0.864	0.902	1.043	1.062	1.137	1.149	1.203	1.209	1.250	1.254
.2	0.842	0.890	1.025	1.057	1.128	1.148	1.196	1.208	1.247	1.254
.3	0.817	0.885	1.003	1.053	1.113	1.147	1.183	1.206	1.241	1.251
.4	0.786	0.878	0.975	1.050	1.090	1.145	1.167	1.206	1.235	1.254
.5	0.755	0.874	0.944	1.049	1.060	1.143	1.150	1.207	1.219	1.253
.6	0.722	0.867	0.916	1.044	1.033	1.139	1.126	1.205	1.208	1.252
FCR	0.8674		1.0675		1.1386		1.1957		1.2459	
5% Score Test: ACR = 1.6449										
.0	1.313	1.334	1.533	1.541	1.627	1.629	1.675	1.675	1.695	1.695
.1	1.292	1.324	1.521	1.534	1.621	1.626	1.674	1.675	1.696	1.694
.2	1.271	1.313	1.507	1.530	1.615	1.626	1.671	1.674	1.697	1.694
.3	1.243	1.306	1.488	1.529	1.607	1.627	1.663	1.670	1.697	1.694
.4	1.208	1.299	1.460	1.524	1.586	1.624	1.654	1.673	1.699	1.697
.5	1.171	1.294	1.428	1.525	1.559	1.623	1.645	1.675	1.691	1.695
.6	1.132	1.287	1.397	1.518	1.533	1.621	1.627	1.674	1.687	1.697
FCR	1.2949		1.5596		1.6069		1.6731		1.6915	
1% Score Test: ACR = 2.3263										
.0	2.325	2.321	2.673	2.664	2.741	2.732	2.717	2.708	2.626	2.622
.1	2.311	2.318	2.671	2.660	2.753	2.737	2.734	2.715	2.643	2.627
.2	2.293	2.309	2.679	2.666	2.771	2.746	2.749	2.714	2.658	2.625
.3	2.261	2.297	2.669	2.664	2.785	2.746	2.763	2.709	2.682	2.623
.4	2.221	2.293	2.652	2.662	2.785	2.739	2.786	2.715	2.715	2.633
.5	2.170	2.286	2.625	2.665	2.778	2.742	2.809	2.724	2.743	2.630
.6	2.117	2.276	2.591	2.651	2.764	2.745	2.821	2.730	2.771	2.632
FCR	2.3155		2.6806		2.7381		2.6782		2.6316	

Note: ACR = asymptotic critical value;
FCR = Finite sample critical value based on 50,000 Monte Carlo Samples.

Table 1b. Rejection Rates of Score Test for IG Distribution: $m = \omega = 2$

Uncensored Data

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.0470	.0238	.0106	.0928	.0462	.0098	.0772	.0388	.0100
	.1	.0872	.0556	.0198	.1486	.0858	.0228	.1294	.0764	.0224
	.2	.1130	.0766	.0390	.1958	.1152	.0418	.1704	.0990	.0392
	.3	.1634	.1114	.0622	.2536	.1684	.0698	.2246	.1480	.0638
	.4	.2012	.1482	.0840	.3154	.2240	.1054	.2754	.1918	.0850
	.5	.2486	.1816	.1060	.3802	.2786	.1402	.3308	.2344	.1070
	.6	.2614	.1968	.1086	.4052	.3054	.1558	.3476	.2500	.1126
50	.0	.0770	.0474	.0158	.1020	.0552	.0108	.0964	.0536	.0112
	.1	.1474	.1020	.0508	.1928	.1170	.0362	.1824	.1120	.0390
	.2	.2662	.1964	.1254	.3252	.2218	.1020	.3146	.2120	.1040
	.3	.3798	.3122	.2080	.4458	.3434	.1794	.4304	.3332	.1798
	.4	.4824	.4096	.3056	.5526	.4430	.2678	.5372	.4318	.2676
	.5	.5828	.5128	.4044	.6524	.5476	.3606	.6362	.5356	.3586
	.6	.6320	.5638	.4470	.7086	.6054	.4084	.6876	.5868	.4048
100	.0	.0824	.0448	.0172	.1040	.0476	.0086	.0998	.0476	.0096
	.1	.2262	.1586	.0866	.2534	.1616	.0590	.2488	.1594	.0632
	.2	.4104	.3272	.2146	.4510	.3384	.1684	.4462	.3320	.1718
	.3	.5734	.4930	.3734	.6124	.5006	.3104	.6054	.4992	.3160
	.4	.7126	.6418	.5292	.7464	.6506	.4700	.7422	.6470	.4748
	.5	.8114	.7562	.6558	.8418	.7672	.6002	.8370	.7620	.6018
	.6	.8698	.8266	.7476	.8944	.8356	.6904	.8892	.8302	.6934
200	.0	.0882	.0484	.0152	.1004	.0470	.0080	.0992	.0470	.0088
	.1	.3172	.2358	.1280	.3384	.2330	.0926	.3362	.2338	.0944
	.2	.5978	.5048	.3596	.6200	.4980	.2948	.6162	.4984	.2998
	.3	.8144	.7548	.6342	.8304	.7514	.5624	.8274	.7516	.5710
	.4	.9152	.8814	.8066	.9236	.8802	.7540	.9218	.8798	.7600
	.5	.9676	.9500	.8990	.9724	.9490	.8628	.9710	.9492	.8674
	.6	.9866	.9776	.9558	.9890	.9774	.9312	.9880	.9764	.9330
500	.0	.1018	.0576	.0178	.1060	.0534	.0114	.1050	.0536	.0112
	.1	.5156	.4082	.2428	.5254	.3950	.1908	.5252	.3944	.1958
	.2	.8774	.8154	.6878	.8830	.8050	.6246	.8822	.8054	.6326
	.3	.9818	.9670	.9302	.9818	.9654	.8990	.9820	.9650	.9028
	.4	.9986	.9976	.9918	.9990	.9970	.9824	.9990	.9970	.9828
	.5	.9998	.9992	.9980	.9998	.9992	.9968	.9998	.9992	.9966
	.6	.9994	.9994	.9992	1.0000	.9998	.9992	.9996	.9994	.9990

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 2a. Bootstrap Critical Values for Ducharme's GOF Test: $m = \omega = 2$

Uncensored Data

	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
10% Score Test: ACR = 4.6052										
.0	2.719	2.699	3.194	3.207	3.485	3.495	3.733	3.739	4.004	4.009
.1	2.716	2.695	3.179	3.199	3.465	3.489	3.711	3.730	3.986	4.001
.2	2.707	2.690	3.163	3.194	3.452	3.488	3.699	3.734	3.971	4.004
.3	2.700	2.688	3.148	3.193	3.428	3.483	3.672	3.727	3.956	4.006
.4	2.686	2.681	3.127	3.191	3.402	3.481	3.655	3.732	3.931	4.000
.5	2.681	2.681	3.108	3.186	3.381	3.479	3.632	3.729	3.911	4.001
.6	2.672	2.681	3.094	3.181	3.363	3.475	3.613	3.725	3.888	3.994
FCR	2.7253		3.1791		3.4344		3.7609		3.9763	
5% Score Test: ACR = 5.9915										
.0	3.903	3.900	4.649	4.669	5.104	5.118	5.453	5.463	5.776	5.779
.1	3.897	3.896	4.623	4.656	5.071	5.098	5.426	5.449	5.768	5.778
.2	3.881	3.885	4.604	4.652	5.061	5.107	5.421	5.459	5.762	5.776
.3	3.873	3.884	4.582	4.650	5.029	5.094	5.390	5.448	5.757	5.787
.4	3.852	3.875	4.562	4.655	5.000	5.094	5.376	5.462	5.732	5.777
.5	3.842	3.874	4.536	4.648	4.970	5.091	5.350	5.451	5.717	5.778
.6	3.827	3.875	4.511	4.634	4.952	5.095	5.328	5.446	5.703	5.766
FCR	3.9236		4.6509		5.0010		5.5309		5.7420	
1% Score Test: ACR = 9.2103										
.0	7.652	7.856	10.530	10.602	12.445	12.440	13.479	13.447	13.552	13.520
.1	7.623	7.866	10.428	10.531	12.394	12.374	13.565	13.489	13.639	13.505
.2	7.575	7.829	10.383	10.514	12.494	12.483	13.710	13.565	13.812	13.561
.3	7.553	7.841	10.332	10.503	12.458	12.440	13.743	13.528	13.943	13.544
.4	7.533	7.834	10.377	10.579	12.427	12.410	13.885	13.593	14.102	13.556
.5	7.513	7.824	10.344	10.547	12.499	12.482	13.849	13.521	14.251	13.585
.6	7.527	7.821	10.301	10.515	12.547	12.482	13.985	13.605	14.350	13.588
FCR	7.6164		10.1464		11.2649		14.0669		13.0196	

Note: ACR = asymptotic critical value;

FCR = Finite sample critical value based on 50,000 Monte Carlo Samples.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 2b. Empirical Rejection Rates for Ducharme's GOF Test: $m = \omega = 2$
Uncensored Data

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.0322	.0144	.0038	.0972	.0470	.0080	.1024	.0480	.0064
	.1	.0342	.0176	.0052	.1152	.0568	.0104	.1200	.0588	.0078
	.2	.0452	.0208	.0036	.1284	.0660	.0094	.1306	.0670	.0054
	.3	.0670	.0346	.0090	.1606	.0946	.0188	.1642	.0950	.0106
	.4	.0862	.0442	.0080	.2050	.1202	.0214	.2060	.1160	.0114
	.5	.1222	.0642	.0070	.2522	.1648	.0276	.2520	.1604	.0144
	.6	.1352	.0692	.0070	.2910	.1908	.0272	.2872	.1838	.0120
50	.0	.0500	.0272	.0114	.1012	.0484	.0110	.1004	.0478	.0106
	.1	.0584	.0318	.0106	.1196	.0568	.0094	.1174	.0538	.0084
	.2	.1070	.0644	.0224	.1876	.1104	.0160	.1830	.1054	.0146
	.3	.1906	.1236	.0570	.2832	.1912	.0466	.2776	.1844	.0392
	.4	.2738	.1978	.1012	.3828	.2764	.0856	.3732	.2688	.0774
	.5	.3720	.2892	.1572	.4942	.3774	.1360	.4894	.3668	.1252
	.6	.4372	.3450	.1978	.5554	.4436	.1672	.5486	.4320	.1564
100	.0	.0604	.0360	.0154	.1034	.0494	.0112	.1026	.0492	.0110
	.1	.0866	.0512	.0160	.1404	.0722	.0098	.1366	.0718	.0104
	.2	.1846	.1258	.0498	.2642	.1642	.0236	.2596	.1602	.0248
	.3	.3202	.2394	.1316	.4076	.2974	.0842	.4010	.2910	.0844
	.4	.4908	.4016	.2584	.5844	.4634	.1792	.5794	.4596	.1788
	.5	.6280	.5464	.3968	.7114	.6042	.2924	.7092	.6002	.2912
	.6	.7312	.6640	.5032	.8020	.7118	.3870	.7992	.7070	.3872
200	.0	.0676	.0390	.0150	.0994	.0486	.0076	.1002	.0484	.0078
	.1	.1172	.0756	.0274	.1670	.0890	.0118	.1636	.0872	.0122
	.2	.3066	.2244	.1120	.3846	.2562	.0496	.3798	.2530	.0530
	.3	.5776	.4842	.3080	.6522	.5234	.1876	.6478	.5180	.1964
	.4	.7780	.7070	.5486	.8264	.7402	.3914	.8254	.7386	.3996
	.5	.9028	.8550	.7386	.9314	.8770	.5812	.9316	.8752	.5916
	.6	.9566	.9330	.8562	.9706	.9444	.7306	.9694	.9428	.7388
500	.0	.0808	.0522	.0238	.1058	.0568	.0112	.1058	.0570	.0114
	.1	.2060	.1380	.0540	.2572	.1456	.0180	.2546	.1462	.0198
	.2	.6310	.5318	.3370	.6794	.5414	.1742	.6760	.5426	.1826
	.3	.9072	.8630	.7434	.9302	.8728	.5528	.9278	.8728	.5686
	.4	.9894	.9808	.9416	.9936	.9822	.8542	.9932	.9816	.8616
	.5	.9982	.9970	.9886	.9986	.9972	.9640	.9986	.9970	.9670
	.6	.9998	.9998	.9982	.9998	.9998	.9916	.9998	.9998	.9918

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Table 3a. Bootstrap Critical Values of Score Test for IG Distribution: $m = \omega = 2$
Censoring Time = 90th Percentile of the True Distribution.

	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
	10% Score Test: ACR = 1.2816									
.0	1.747	1.713	1.553	1.567	1.464	1.476	1.401	1.407	1.347	1.349
.1	1.746	1.725	1.543	1.562	1.457	1.473	1.396	1.406	1.347	1.352
.2	1.727	1.717	1.534	1.561	1.452	1.473	1.391	1.404	1.344	1.351
.3	1.714	1.719	1.519	1.553	1.442	1.469	1.387	1.400	1.344	1.350
.4	1.709	1.720	1.524	1.566	1.436	1.467	1.387	1.405	1.342	1.352
.5	1.701	1.725	1.514	1.568	1.432	1.467	1.383	1.404	1.340	1.350
.6	1.696	1.737	1.503	1.556	1.429	1.469	1.380	1.402	1.337	1.349
FCR	1.7521		1.5652		1.4438		1.4090		1.3496	
	5% Score Test: ACR = 1.6449									
.0	2.214	2.202	1.998	2.021	1.888	1.902	1.806	1.814	1.734	1.737
.1	2.210	2.211	1.986	2.013	1.880	1.900	1.802	1.815	1.734	1.740
.2	2.194	2.202	1.977	2.013	1.873	1.901	1.795	1.809	1.731	1.740
.3	2.184	2.206	1.960	2.005	1.859	1.891	1.788	1.806	1.729	1.738
.4	2.180	2.206	1.961	2.019	1.854	1.894	1.788	1.812	1.728	1.739
.5	2.171	2.208	1.951	2.017	1.849	1.893	1.781	1.809	1.724	1.737
.6	2.170	2.221	1.941	2.006	1.842	1.894	1.777	1.808	1.718	1.734
FCR	2.2267		2.0176		1.8620		1.8299		1.7312	
	1% Score Test: ACR = 2.3263									
.0	3.030	3.030	2.843	2.879	2.707	2.727	2.578	2.593	2.478	2.482
.1	3.034	3.041	2.831	2.874	2.693	2.719	2.585	2.604	2.480	2.485
.2	3.019	3.030	2.822	2.865	2.682	2.718	2.569	2.592	2.468	2.481
.3	3.020	3.031	2.798	2.854	2.672	2.724	2.562	2.591	2.463	2.474
.4	3.025	3.031	2.805	2.880	2.658	2.715	2.557	2.590	2.464	2.485
.5	3.028	3.033	2.794	2.873	2.654	2.723	2.552	2.592	2.458	2.485
.6	3.038	3.047	2.776	2.860	2.639	2.714	2.544	2.589	2.450	2.472
FCR	3.0625		2.8928		2.6839		2.6186		2.4657	

Note: ACR = asymptotic critical value;
FCR = Finite sample critical value based on 50,000 Monte Carlo Samples.

Table 3b. Rejection Rates of Score Test for IG Distribution: $m = \omega = 2$
Censoring Time = 90th Percentile of the True Distribution.

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.1725	.1085	.0455	.0850	.0425	.0030	.1200	.0565	.0050
	.1	.2215	.1540	.0535	.1175	.0475	.0055	.1415	.0610	.0065
	.2	.2435	.1590	.0615	.1315	.0645	.0120	.1450	.0755	.0145
	.3	.2620	.1705	.0720	.1400	.0750	.0135	.1550	.0790	.0165
	.4	.3100	.2055	.0865	.1780	.0905	.0155	.1755	.0920	.0165
	.5	.3290	.2375	.1055	.2140	.1160	.0215	.2055	.1145	.0225
	.6	.3690	.2575	.1150	.2310	.1285	.0205	.2190	.1150	.0220
50	.0	.1490	.0915	.0270	.0970	.0455	.0070	.0855	.0440	.0055
	.1	.1795	.1130	.0370	.1255	.0620	.0125	.1130	.0550	.0095
	.2	.2365	.1490	.0550	.1685	.0890	.0175	.1490	.0710	.0140
	.3	.3085	.2100	.0880	.2305	.1365	.0360	.2110	.1210	.0280
	.4	.3960	.2720	.1070	.3035	.1715	.0455	.2815	.1415	.0335
	.5	.4435	.3235	.1380	.3550	.2200	.0590	.3305	.1875	.0405
	.6	.4615	.3280	.1460	.3725	.2285	.0690	.3480	.1950	.0485
100	.0	.1300	.0800	.0225	.1050	.0495	.0120	.0985	.0405	.0090
	.1	.2135	.1320	.0470	.1685	.0890	.0250	.1560	.0775	.0165
	.2	.2810	.1855	.0725	.2320	.1300	.0350	.2205	.1175	.0295
	.3	.3900	.2890	.1245	.3455	.2235	.0745	.3350	.2075	.0550
	.4	.4715	.3500	.1600	.4145	.2815	.1005	.4035	.2620	.0770
	.5	.5805	.4500	.2290	.5320	.3695	.1495	.5220	.3510	.1305
	.6	.6505	.5200	.2855	.5995	.4490	.1995	.5890	.4360	.1665
200	.0	.1100	.0645	.0215	.0930	.0500	.0130	.0885	.0495	.0100
	.1	.2520	.1560	.0495	.2130	.1210	.0330	.2110	.1140	.0280
	.2	.3600	.2460	.0995	.3220	.2065	.0650	.3150	.2020	.0580
	.3	.5075	.3875	.1735	.4670	.3395	.1200	.4640	.3300	.1130
	.4	.6510	.5170	.2915	.6175	.4700	.2165	.6155	.4640	.2005
	.5	.7695	.6480	.4135	.7410	.6040	.3320	.7320	.5950	.3190
	.6	.8495	.7575	.5225	.8315	.7255	.4460	.8250	.7155	.4245
500	.0	.1170	.0615	.0160	.1050	.0505	.0110	.1015	.0495	.0100
	.1	.3115	.1920	.0585	.2905	.1625	.0435	.2880	.1590	.0415
	.2	.5390	.4020	.1925	.5170	.3730	.1600	.5130	.3690	.1515
	.3	.7660	.6465	.3950	.7510	.6160	.3450	.7505	.6130	.3435
	.4	.9090	.8360	.6195	.8995	.8160	.5675	.8995	.8160	.5650
	.5	.9745	.9330	.7865	.9710	.9205	.7500	.9710	.9195	.7465
	.6	.9890	.9795	.9040	.9885	.9755	.8835	.9885	.9755	.8815

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Table 4a. Bootstrap Critical Values of Score Test for IG Distribution: $m = \omega = 2$
Censoring Time $= 80$th Percentile of the True Distribution.

	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
	10% Score Test: ACR = 1.2816									
.0	1.995	1.960	1.709	1.728	1.569	1.581	1.471	1.479	1.391	1.395
.1	1.977	1.955	1.704	1.729	1.561	1.584	1.468	1.480	1.385	1.394
.2	1.964	2.005	1.690	1.736	1.555	1.579	1.458	1.475	1.384	1.393
.3	1.964	2.011	1.684	1.732	1.552	1.577	1.457	1.477	1.383	1.394
.4	1.955	2.011	1.682	1.731	1.541	1.577	1.455	1.476	1.383	1.395
.5	1.940	1.998	1.672	1.735	1.540	1.577	1.448	1.476	1.379	1.394
.6	1.949	2.020	1.665	1.732	1.538	1.580	1.447	1.475	1.379	1.393
FCR	1.9854		1.7080		1.5449		1.4758		1.3920	
	5% Score Test: ACR = 1.6449									
.0	2.489	2.479	2.191	2.221	2.020	2.040	1.900	1.908	1.792	1.796
.1	2.473	2.469	2.183	2.221	2.013	2.041	1.890	1.911	1.785	1.795
.2	2.460	2.504	2.171	2.224	2.003	2.035	1.882	1.904	1.786	1.796
.3	2.460	2.509	2.161	2.218	1.998	2.035	1.879	1.903	1.780	1.793
.4	2.456	2.508	2.155	2.218	1.989	2.030	1.876	1.902	1.780	1.795
.5	2.440	2.494	2.144	2.219	1.986	2.032	1.868	1.904	1.775	1.791
.6	2.452	2.520	2.139	2.221	1.982	2.038	1.867	1.903	1.775	1.795
FCR	2.4886		2.1763		2.0015		1.8997		1.7865	
	1% Score Test: ACR = 2.3263									
.0	3.319	3.317	3.115	3.161	2.902	2.935	2.728	2.743	2.568	2.567
.1	3.314	3.312	3.098	3.153	2.890	2.938	2.717	2.746	2.560	2.566
.2	3.307	3.321	3.083	3.152	2.881	2.920	2.706	2.733	2.551	2.568
.3	3.315	3.333	3.073	3.144	2.869	2.927	2.701	2.735	2.551	2.575
.4	3.322	3.335	3.062	3.139	2.853	2.923	2.695	2.733	2.544	2.572
.5	3.317	3.322	3.047	3.147	2.856	2.926	2.687	2.733	2.540	2.565
.6	3.343	3.360	3.049	3.141	2.843	2.931	2.683	2.735	2.538	2.569
FCR	3.3287		3.1078		2.9082		2.7457		2.5591	

Note: ACR = asymptotic critical value;
FCR = Finite sample critical value based on 50,000 Monte Carlo Samples.

Table 4b. Rejection Rates of Score Test for IG Distribution: $m = \omega = 2$
Censoring Time = 80th Percentile of the True Distribution.

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.2145	.1450	.0540	.0760	.0375	.0040	.1105	.0450	.0070
	.1	.2320	.1655	.0625	.0950	.0365	.0040	.1210	.0520	.0055
	.2	.2635	.1795	.0780	.1125	.0570	.0120	.1005	.0520	.0115
	.3	.3020	.2095	.0980	.1400	.0675	.0120	.1255	.0590	.0130
	.4	.3080	.2205	.0960	.1450	.0675	.0130	.1295	.0615	.0130
	.5	.3005	.2150	.1045	.1485	.0765	.0120	.1380	.0700	.0100
	.6	.3695	.2595	.1230	.1860	.0950	.0225	.1590	.0805	.0205
50	.0	.1640	.1050	.0445	.0915	.0495	.0075	.0825	.0515	.0065
	.1	.1975	.1265	.0520	.1130	.0595	.0150	.1005	.0525	.0085
	.2	.2380	.1520	.0630	.1340	.0780	.0175	.1155	.0595	.0130
	.3	.2700	.1885	.0725	.1785	.0910	.0210	.1530	.0715	.0130
	.4	.3070	.2075	.0835	.1985	.1060	.0265	.1755	.0830	.0160
	.5	.3545	.2440	.1140	.2335	.1370	.0425	.2075	.1145	.0280
	.6	.3965	.2775	.1360	.2675	.1620	.0475	.2385	.1355	.0335
100	.0	.1530	.0930	.0275	.1020	.0555	.0080	.0955	.0395	.0055
	.1	.1920	.1190	.0360	.1285	.0625	.0120	.1195	.0510	.0070
	.2	.2230	.1495	.0585	.1665	.0980	.0280	.1520	.0770	.0195
	.3	.3025	.2030	.0810	.2270	.1235	.0320	.2135	.1115	.0235
	.4	.3725	.2705	.1180	.2960	.1805	.0485	.2860	.1570	.0365
	.5	.4330	.3130	.1355	.3495	.2110	.0650	.3315	.1845	.0490
	.6	.5200	.3890	.1950	.4285	.2875	.1050	.4080	.2680	.0720
200	.0	.1280	.0795	.0265	.0985	.0555	.0095	.0955	.0520	.0085
	.1	.1970	.1215	.0440	.1500	.0815	.0220	.1475	.0775	.0145
	.2	.2810	.1845	.0740	.2300	.1360	.0410	.2235	.1200	.0315
	.3	.3965	.2750	.1280	.3350	.2170	.0765	.3320	.2090	.0630
	.4	.4760	.3450	.1590	.4150	.2700	.1040	.4040	.2640	.0900
	.5	.5720	.4380	.2330	.5145	.3635	.1445	.5060	.3510	.1315
	.6	.6600	.5385	.3055	.6045	.4510	.2100	.6045	.4405	.1925
500	.0	.1135	.0540	.0165	.0950	.0425	.0105	.0935	.0405	.0125
	.1	.2170	.1350	.0415	.1890	.1070	.0230	.1845	.1070	.0240
	.2	.4000	.2825	.1315	.3660	.2400	.0990	.3605	.2385	.0900
	.3	.5460	.4240	.2145	.5135	.3815	.1555	.5150	.3790	.1460
	.4	.7015	.5610	.3125	.6660	.5055	.2460	.6640	.5070	.2335
	.5	.8110	.7055	.4560	.7835	.6710	.3765	.7755	.6620	.3675
	.6	.8795	.8070	.5955	.8625	.7705	.5265	.8635	.7700	.5120

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Table 5a. Bootstrap Critical Values of Score Test for IG Distribution: $m = \omega = 2$
$C \sim$ Uniform$(.8q_{.8}, 1.2q_{.8})$, $q_{.8} = .8$-quantile of IGM.

	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
	10% Score Test: ACR = 1.2816									
.0	1.913	1.880	1.646	1.656	1.522	1.525	1.435	1.444	1.374	1.376
.1	1.928	1.906	1.629	1.638	1.518	1.530	1.436	1.442	1.369	1.374
.2	1.862	1.856	1.624	1.637	1.509	1.520	1.432	1.441	1.367	1.372
.3	1.843	1.843	1.647	1.665	1.511	1.526	1.432	1.445	1.367	1.371
.4	1.903	1.908	1.622	1.642	1.504	1.528	1.432	1.441	1.369	1.372
.5	1.879	1.887	1.612	1.637	1.498	1.518	1.429	1.440	1.367	1.372
.6	1.851	1.869	1.616	1.639	1.501	1.520	1.428	1.441	1.364	1.372
FCR	1.8642		1.6440		1.5070		1.4324		1.3626	
	5% Score Test: ACR = 1.6449									
.0	2.374	2.340	2.102	2.114	1.952	1.960	1.848	1.857	1.769	1.768
.1	2.394	2.367	2.077	2.085	1.951	1.965	1.848	1.855	1.759	1.767
.2	2.322	2.311	2.078	2.086	1.940	1.953	1.840	1.854	1.759	1.765
.3	2.303	2.295	2.104	2.118	1.943	1.959	1.844	1.859	1.758	1.766
.4	2.374	2.362	2.074	2.095	1.933	1.959	1.841	1.852	1.758	1.765
.5	2.349	2.346	2.063	2.090	1.927	1.950	1.839	1.855	1.759	1.766
.6	2.320	2.321	2.064	2.090	1.932	1.953	1.836	1.852	1.755	1.765
FCR	2.3253		2.0896		1.9393		1.8426		1.7496	
	1% Score Test: ACR = 2.3263									
.0	3.132	3.086	2.956	2.963	2.774	2.787	2.647	2.655	2.518	2.526
.1	3.152	3.108	2.921	2.928	2.786	2.795	2.643	2.652	2.511	2.521
.2	3.090	3.052	2.922	2.922	2.762	2.784	2.629	2.646	2.508	2.518
.3	3.081	3.034	2.962	2.971	2.773	2.782	2.636	2.657	2.507	2.518
.4	3.171	3.111	2.924	2.935	2.750	2.785	2.629	2.641	2.516	2.518
.5	3.134	3.091	2.909	2.926	2.749	2.773	2.628	2.653	2.506	2.520
.6	3.122	3.077	2.912	2.936	2.751	2.781	2.626	2.649	2.496	2.523
FCR	3.1107		2.9114		2.7927		2.6124		2.5295	

Note: ACR = asymptotic critical value;
FCR = Finite sample critical value based on 50,000 Monte Carlo Samples.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 5b. Rejection Rates of Score Test for IG Distribution: $m = \omega = 2$
$C \sim$ U$(.8q_{.8}, 1.2q_{.8})$, $q_{.8} = .8$-quantile of IGM; Treated as Known

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.2050	.1340	.0575	.0980	.0465	.0090	.1115	.0565	.0150
	.1	.2195	.1505	.0590	.1010	.0475	.0045	.1080	.0550	.0105
	.2	.2550	.1745	.0710	.1305	.0640	.0110	.1355	.0680	.0170
	.3	.2660	.1835	.0730	.1375	.0635	.0145	.1405	.0715	.0180
	.4	.2965	.2070	.0985	.1510	.0780	.0170	.1555	.0830	.0210
	.5	.3060	.2205	.0950	.1600	.0830	.0190	.1660	.0910	.0240
	.6	.3280	.2315	.1030	.1785	.0925	.0165	.1770	.0925	.0230
50	.0	.1655	.1070	.0360	.1055	.0470	.0070	.1000	.0430	.0085
	.1	.1895	.1235	.0455	.1215	.0645	.0125	.1160	.0615	.0130
	.2	.2430	.1625	.0620	.1650	.0820	.0200	.1555	.0845	.0205
	.3	.2830	.1885	.0785	.1840	.1090	.0250	.1780	.1000	.0185
	.4	.3065	.2090	.0780	.2150	.1130	.0235	.2040	.1060	.0245
	.5	.3330	.2375	.0965	.2410	.1370	.0315	.2375	.1210	.0275
	.6	.3840	.2835	.1325	.2915	.1765	.0475	.2775	.1670	.0415
100	.0	.1340	.0785	.0230	.0935	.0450	.0090	.0890	.0425	.0080
	.1	.1960	.1190	.0290	.1400	.0670	.0115	.1365	.0595	.0100
	.2	.2490	.1535	.0540	.1880	.0990	.0230	.1785	.0940	.0200
	.3	.3190	.2090	.0770	.2520	.1385	.0340	.2420	.1305	.0345
	.4	.3895	.2680	.1260	.3095	.1945	.0620	.2995	.1865	.0590
	.5	.4425	.3095	.1370	.3585	.2290	.0725	.3510	.2130	.0610
	.6	.4785	.3550	.1730	.3975	.2650	.0945	.3945	.2610	.0870
200	.0	.1195	.0635	.0175	.0930	.0450	.0090	.0870	.0450	.0090
	.1	.1955	.1215	.0405	.1645	.0865	.0245	.1610	.0870	.0215
	.2	.2875	.1890	.0730	.2405	.1440	.0425	.2345	.1375	.0400
	.3	.3775	.2620	.1040	.3285	.2085	.0645	.3265	.1950	.0590
	.4	.4930	.3540	.1565	.4285	.2895	.1020	.4275	.2775	.0985
	.5	.5830	.4450	.2330	.5230	.3775	.1530	.5205	.3730	.1425
	.6	.6695	.5455	.3075	.6210	.4715	.2240	.6140	.4645	.2170
500	.0	.1100	.0580	.0160	.0945	.0470	.0070	.0915	.0485	.0095
	.1	.2425	.1415	.0420	.2140	.1130	.0305	.2140	.1140	.0300
	.2	.3940	.2645	.0955	.3625	.2250	.0685	.3555	.2260	.0715
	.3	.5670	.4325	.2235	.5350	.3980	.1775	.5355	.3890	.1735
	.4	.7245	.5875	.3205	.6880	.5365	.2690	.6940	.5380	.2660
	.5	.8450	.7420	.5145	.8160	.7065	.4540	.8185	.7050	.4450
	.6	.8915	.8145	.6025	.8750	.7895	.5465	.8745	.7825	.5390

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 5c. Rejection Rates of Score Test for IG Distribution: $m = \omega = 2$
$C \sim$ `U`$(.8q_{.8}, 1.2q_{.8})$, $q_{.8} = .8$-quantile of IGM; Treated as Unknown

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
20	.0	.2075	.1355	.0470	.0910	.0420	.0075	.1070	.0545	.0115
	.1	.2265	.1465	.0615	.0985	.0490	.0085	.1135	.0595	.0130
	.2	.2415	.1625	.0680	.1090	.0560	.0125	.1325	.0735	.0175
	.3	.2795	.1910	.0850	.1390	.0685	.0165	.1550	.0915	.0245
	.4	.3000	.2195	.1005	.1450	.0785	.0135	.1695	.0935	.0205
	.5	.3230	.2385	.1125	.1625	.0855	.0105	.1855	.1035	.0225
	.6	.3175	.2335	.1040	.1540	.0810	.0175	.1785	.1035	.0325
50	.0	.1615	.1020	.0395	.0975	.0505	.0080	.1070	.0560	.0145
	.1	.1905	.1285	.0480	.1280	.0690	.0120	.1395	.0805	.0170
	.2	.2455	.1580	.0565	.1505	.0750	.0170	.1670	.0890	.0205
	.3	.2600	.1795	.0680	.1695	.0900	.0150	.1835	.0990	.0270
	.4	.3080	.2115	.0955	.2000	.1225	.0305	.2155	.1365	.0430
	.5	.3475	.2470	.1085	.2405	.1370	.0365	.2570	.1495	.0475
	.6	.3260	.2330	.0975	.2175	.1165	.0345	.2300	.1300	.0415
100	.0	.1405	.0780	.0220	.0930	.0445	.0060	.0990	.0540	.0085
	.1	.2030	.1180	.0425	.1380	.0740	.0130	.1470	.0795	.0180
	.2	.2490	.1605	.0650	.1775	.1020	.0250	.1905	.1120	.0345
	.3	.3000	.1975	.0755	.2230	.1250	.0320	.2340	.1270	.0350
	.4	.3455	.2270	.0980	.2650	.1480	.0495	.2690	.1595	.0520
	.5	.3880	.2840	.1300	.3090	.1945	.0590	.3170	.2100	.0635
	.6	.4105	.2955	.1340	.3180	.1985	.0570	.3320	.2105	.0675
200	.0	.1285	.0700	.0160	.1035	.0500	.0065	.1060	.0505	.0085
	.1	.2000	.1225	.0315	.1640	.0815	.0165	.1685	.0855	.0150
	.2	.2760	.1795	.0610	.2320	.1285	.0360	.2355	.1375	.0400
	.3	.3465	.2390	.0985	.2945	.1815	.0540	.2980	.1840	.0620
	.4	.4405	.3250	.1350	.3845	.2480	.0820	.3915	.2520	.0860
	.5	.5245	.3895	.1815	.4535	.3115	.1125	.4550	.3185	.1185
	.6	.5535	.4205	.2160	.4885	.3580	.1240	.4980	.3610	.1350
500	.0	.1275	.0735	.0175	.1160	.0570	.0095	.1145	.0570	.0095
	.1	.2380	.1445	.0425	.2145	.1175	.0315	.2100	.1240	.0320
	.2	.3670	.2430	.0990	.3320	.2145	.0615	.3330	.2170	.0710
	.3	.5110	.3650	.1535	.4720	.3130	.1105	.4715	.3105	.1195
	.4	.6125	.4770	.2360	.5765	.4285	.1785	.5770	.4325	.1915
	.5	.7260	.5925	.3625	.6915	.5420	.2955	.6955	.5480	.3070
	.6	.8260	.7120	.4635	.7950	.6760	.3765	.7965	.6810	.3840

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Table 6a. Finite Sample Critical Values of Score Test for IG Regression
`Complete Data,` $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

n	1%	5%	10%	90%	95%	99%	Mean	StD
20	-3.936	-3.525	-3.203	1.320	1.944	2.924	-1.129	1.697
50	-4.663	-3.725	-3.125	1.232	1.708	2.532	-0.877	1.655
100	-4.511	-3.350	-2.705	1.185	1.587	2.326	-0.654	1.511
200	-4.266	-2.996	-2.352	1.131	1.499	2.139	-0.508	1.377
500	-3.659	-2.543	-1.973	1.150	1.470	2.054	-0.327	1.230
1000	-3.328	-2.308	-1.798	1.143	1.454	2.008	-0.255	1.154
2000	-3.054	-2.108	-1.634	1.150	1.471	2.030	-0.185	1.093
5000	-2.788	-1.925	-1.496	1.178	1.494	2.089	-0.118	1.046
10000	-2.711	-1.862	-1.441	1.188	1.515	2.117	-0.098	1.031
ACR	-2.236	-1.645	-1.282	1.282	1.645	2.236	0.000	1.000

Note: Each set of finite sample critical values are based on $M = 50,000$.

Table 6b. Bootstrap Critical Values of Score Test for IG Regression
`Complete Data,` $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

	$n = 50$		$n = 100$		$n = 200$		$n = 500$	
v	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE	rMLE	uMLE
10% Score Test: ACR = 1.2816								
0.00	1.261	1.256	1.180	1.180	1.149	1.151	1.133	1.134
0.05	1.220	1.216	1.179	1.181	1.147	1.153	1.124	1.127
0.10	1.233	1.235	1.169	1.177	1.133	1.140	1.117	1.123
0.15	1.214	1.221	1.167	1.177	1.130	1.140	1.111	1.119
0.20	1.221	1.233	1.126	1.141	1.112	1.125	1.100	1.111
0.25	1.146	1.165	1.128	1.145	1.102	1.117	1.090	1.101
FCR	1.2319		1.1850		1.1310		1.1495	
5% Score Test: ACR = 1.6449								
0.00	1.727	1.722	1.576	1.577	1.505	1.505	1.459	1.460
0.05	1.682	1.682	1.584	1.586	1.506	1.510	1.447	1.449
0.10	1.702	1.704	1.573	1.578	1.486	1.494	1.443	1.448
0.15	1.682	1.688	1.568	1.578	1.485	1.497	1.434	1.441
0.20	1.691	1.702	1.537	1.549	1.466	1.478	1.418	1.431
0.25	1.642	1.656	1.530	1.544	1.462	1.475	1.407	1.420
FCR	1.7075		1.5868		1.4985		1.4700	
1% Score Test: ACR = 2.3263								
0.00	2.547	2.543	2.276	2.273	2.129	2.131	2.034	2.034
0.05	2.485	2.488	2.293	2.296	2.135	2.139	2.019	2.022
0.10	2.516	2.521	2.275	2.280	2.112	2.119	2.017	2.024
0.15	2.493	2.499	2.274	2.281	2.110	2.119	2.002	2.012
0.20	2.510	2.515	2.257	2.264	2.087	2.095	1.980	1.990
0.25	2.501	2.504	2.235	2.242	2.092	2.101	1.961	1.973
FCR	2.5322		2.3257		2.1389		2.0541	

Note: ACR = asymptotic critical value;
FCR = Finite sample critical value based on $M = 50,000$.

Table 6c. Empirical Rejection Rates of Score Test for IG Regression
Complete Data, $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

		Based on ACR			Based on BCRr			Based on BCRu		
n	v	10%	5%	1%	10%	5%	1%	10%	5%	1%
50	.00	.1025	.0570	.0180	.1055	.0510	.0090	.0985	.0455	.0085
	.05	.1140	.0685	.0225	.1260	.0665	.0155	.1230	.0655	.0150
	.10	.1820	.1245	.0410	.1895	.1160	.0340	.1855	.1115	.0295
	.15	.2570	.1760	.0820	.2670	.1665	.0670	.2605	.1630	.0620
	.20	.3200	.2460	.1350	.3345	.2360	.1145	.3280	.2290	.1090
	.25	.3900	.3155	.1775	.4215	.3165	.1545	.4140	.3100	.1500
100	.00	.0795	.0435	.0105	.0950	.0470	.0130	.0905	.0465	.0115
	.05	.1460	.0835	.0245	.1675	.0925	.0275	.1630	.0895	.0240
	.10	.2790	.1905	.0630	.3030	.2035	.0730	.3020	.1980	.0695
	.15	.3505	.2615	.1175	.3825	.2825	.1300	.3800	.2785	.1265
	.20	.4740	.3855	.2085	.5170	.4070	.2305	.5125	.4010	.2305
	.25	.5890	.4930	.2945	.6340	.5250	.3210	.6310	.5215	.3135
200	.00	.0810	.0395	.0035	.1035	.0555	.0095	.1010	.0555	.0100
	.05	.1960	.1185	.0330	.2295	.1480	.0475	.2250	.1445	.0480
	.10	.3585	.2435	.0965	.4105	.2965	.1320	.4075	.2950	.1330
	.15	.5400	.4145	.1755	.5820	.4700	.2500	.5805	.4630	.2470
	.20	.7025	.5730	.3355	.7485	.6355	.4200	.7490	.6320	.4215
	.25	.8020	.7100	.4770	.8340	.7595	.5670	.8325	.7590	.5585
500	.00	.0665	.0310	.0040	.0900	.0480	.0145	.0920	.0485	.0150
	.05	.2920	.1745	.0360	.3485	.2330	.0865	.3480	.2310	.0835
	.10	.5855	.4475	.1800	.6365	.5160	.3020	.6365	.5140	.3010
	.15	.8195	.7060	.4060	.8540	.7710	.5585	.8540	.7690	.5505
	.20	.9430	.9015	.6670	.9585	.9290	.7965	.9575	.9265	.7915
	.25	.9825	.9535	.8370	.9890	.9740	.9080	.9885	.9730	.9040

Note: ACR = Asymptotic critical value;
BCRr = Bootstrap critical value based on restricted estimates;
BCRu = Bootstrap critical value based on unrestricted estimates.

Table 7a. Finite Sample Critical Values of Score Test for IG Regression
Type II Censored Data, $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

n	1%	5%	10%	90%	95%	99%	Mean	StD
50	-2.197	-1.493	-1.101	1.955	2.421	3.341	0.403	1.196
100	-2.141	-1.490	-1.138	1.669	2.099	2.911	0.249	1.093
200	-2.182	-1.524	-1.170	1.544	1.956	2.745	0.179	1.057
500	-2.225	-1.549	-1.190	1.432	1.811	2.559	0.112	1.025
1000	-2.230	-1.562	-1.210	1.379	1.755	2.474	0.079	1.011
2000	-2.242	-1.578	-1.223	1.350	1.725	2.419	0.053	1.002
5000	-2.285	-1.618	-1.261	1.317	1.697	2.400	0.034	1.006
10000	-2.285	-1.637	-1.270	1.312	1.677	2.365	0.021	1.004
ACR	-2.236	-1.645	-1.282	1.282	1.645	2.236	0.000	1.000

Note: Each set of finite sample critical values are based on $M = 50,000$.

Table 7b. Bootstrap Critical Values of Score Test for IG Regression, based on rMLE
Type II Censored Data, $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

	$n = 50$			$n = 100$			$n = 200$			$n = 500$		
v	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
.00	2.116	2.632	3.675	1.698	2.122	2.924	1.538	1.939	2.702	1.432	1.820	2.551
.05	2.094	2.610	3.691	1.685	2.109	2.908	1.539	1.943	2.706	1.431	1.818	2.545
.10	2.121	2.619	3.664	1.686	2.107	2.902	1.531	1.932	2.689	1.431	1.816	2.551
.15	2.032	2.517	3.497	1.668	2.092	2.883	1.526	1.924	2.683	1.430	1.814	2.544
.20	2.018	2.501	3.473	1.671	2.092	2.882	1.533	1.934	2.693	1.423	1.808	2.534
.25	2.075	2.570	3.589	1.654	2.070	2.858	1.527	1.927	2.689	1.426	1.812	2.536
FCR	1.955	2.421	3.341	1.669	2.099	2.911	1.544	1.956	2.745	1.432	1.811	2.559

Note: ACR = (1.282, 1.645, 2.326); FCR is based on $M = 50,000$.

Table 7c. Empirical Rejection Rates of Score Test for IG Regression
Type II Censored Data, $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

	Based on ACR			Based on BCRr			Based on ACR			Based on BCRr		
v	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
			$n = 50$						$n = 200$			
.00	.2230	.1430	.0475	.0760	.0310	.0025	.1455	.0875	.0215	.1015	.0525	.0080
.05	.2630	.1770	.0730	.1025	.0460	.0090	.1945	.1190	.0355	.1375	.0755	.0200
.10	.2845	.1980	.0815	.1065	.0500	.0080	.2710	.1750	.0635	.2030	.1270	.0320
.15	.3215	.2245	.1000	.1450	.0760	.0130	.3615	.2480	.1115	.2875	.1860	.0580
.20	.3565	.2550	.1060	.1600	.0895	.0195	.4270	.3030	.1350	.3440	.2265	.0890
.25	.3780	.2705	.1260	.1730	.0945	.0255	.5340	.4135	.2115	.4515	.3230	.1295
			$n = 100$						$n = 500$			
.00	.1630	.1020	.0345	.0915	.0460	.0090	.1205	.0670	.0150	.0975	.0485	.0100
.05	.2060	.1320	.0445	.1295	.0680	.0125	.2275	.1455	.0435	.1870	.1120	.0290
.10	.2570	.1680	.0635	.1580	.0915	.0290	.3555	.2440	.0975	.3065	.1935	.0690
.15	.3395	.2360	.0965	.2340	.1430	.0375	.4795	.3590	.1610	.4225	.2970	.1230
.20	.3630	.2555	.1185	.2460	.1550	.0525	.6010	.4740	.2450	.5535	.4040	.1970
.25	.3935	.3065	.1485	.3070	.1965	.0670	.7265	.6175	.3615	.6835	.5505	.3025

Table 8a. Finite Sample Critical Values of Score Test for IG Regression
Random Censored Data, $\gamma = (.5..25, .25)'$, $\beta = (.5..25, .25)'$

n	1%	5%	10%	90%	95%	99%	Mean	StD
50	-2.989	-2.012	-1.527	1.727	2.210	3.075	0.090	1.284
100	-2.583	-1.782	-1.393	1.500	1.918	2.698	0.049	1.127
200	-2.390	-1.689	-1.311	1.423	1.828	2.602	0.053	1.070
500	-2.323	-1.642	-1.271	1.363	1.735	2.466	0.036	1.027
1000	-2.327	-1.640	-1.271	1.348	1.715	2.442	0.033	1.023
2000	-2.295	-1.627	-1.263	1.322	1.683	2.383	0.018	1.007
5000	-2.319	-1.641	-1.268	1.317	1.679	2.371	0.019	1.006
10000	-2.331	-1.644	-1.270	1.294	1.677	2.362	0.010	1.004
ACR	-2.236	-1.645	-1.282	1.282	1.645	2.236	0.000	1.000

Note: Each set of finite sample critical values are based on $M = 50,000$.

Accepted by Journal of Statistical Computation and Simulation on 24/02/2016

Table 8b. Bootstrap Critical Values of Score Test for IG Regression, based on rMLE
Random Censored Data, $\gamma = (.5 .. 25, .25)'$, $\beta = (.5 .. 25, .25)'$

	$n = 50$			$n = 100$			$n = 200$			$n = 500$		
v	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
.00	1.903	2.435	3.518	1.551	1.973	2.778	1.437	1.835	2.593	1.369	1.753	2.483
.05	1.917	2.444	3.531	1.547	1.972	2.779	1.434	1.836	2.594	1.364	1.745	2.469
.10	1.911	2.431	3.509	1.540	1.965	2.773	1.428	1.827	2.581	1.364	1.746	2.474
.15	1.889	2.401	3.448	1.529	1.953	2.755	1.427	1.824	2.582	1.360	1.742	2.470
.20	1.885	2.394	3.429	1.532	1.956	2.758	1.421	1.819	2.577	1.358	1.743	2.471
.25	1.847	2.345	3.358	1.519	1.940	2.739	1.416	1.813	2.570	1.356	1.737	2.459
FCR	1.727	2.210	3.075	1.500	1.918	2.698	1.423	1.828	2.602	1.363	1.735	2.466

Note: ACR = (1.282, 1.645, 2.326); FCR is based on $M = 50,000$.

Table 8c. Empirical Rejection Rates of Score Test for IG Regression
Random Censored Data, $\gamma = (.5 .. 25, .25)'$, $\beta = (.5 .. 25, .25)'$

	Based on ACR			Based on BCRr			Based on ACR			Based on BCRr		
v	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
	$n = 50$						$n = 200$					
.00	.1900	.1195	.0370	.0740	.0300	.0050	.1240	.0710	.0180	.0955	.0475	.0095
.05	.1970	.1310	.0430	.0840	.0310	.0040	.1810	.1040	.0310	.1415	.0705	.0165
.10	.2210	.1425	.0570	.0930	.0440	.0050	.2355	.1460	.0445	.1970	.1100	.0280
.15	.2465	.1695	.0745	.1225	.0590	.0085	.3120	.2090	.0720	.2670	.1630	.0440
.20	.2685	.1770	.0710	.1315	.0590	.0105	.3810	.2490	.1030	.3345	.2005	.0700
.25	.3055	.2160	.0935	.1705	.0855	.0135	.4870	.3580	.1505	.4375	.3000	.1065
	$n = 100$						$n = 500$					
.00	.1445	.0800	.0305	.0920	.0485	.0105	.1190	.0640	.0130	.1000	.0495	.0095
.05	.1810	.1140	.0340	.1250	.0625	.0105	.2030	.1235	.0335	.1840	.1085	.0265
.10	.2105	.1255	.0410	.1490	.0730	.0135	.2955	.1855	.0625	.2670	.1595	.0505
.15	.2735	.1780	.0645	.2010	.1190	.0275	.4685	.3340	.1355	.4375	.2930	.1140
.20	.3080	.2095	.0760	.2360	.1375	.0345	.5875	.4470	.2315	.5545	.4165	.1935
.25	.3540	.2460	.0995	.2805	.1675	.0440	.7090	.5815	.3390	.6870	.5515	.2970

Table 9. Finite Sample Critical Values of Score Test for IG Regression
Type I Censored Data, $\gamma = (.5 .. 25, .25)'$, $\beta = (.5 .. 25, .25)'$

n	1%	5%	10%	90%	95%	99%	Mean	StD
50	-2.799	-1.903	-1.436	1.768	2.255	3.209	0.150	1.272
100	-2.418	-1.696	-1.310	1.549	1.974	2.798	0.094	1.117
200	-2.343	-1.651	-1.273	1.462	1.865	2.656	0.081	1.069
500	-2.310	-1.614	-1.260	1.385	1.766	2.502	0.051	1.030
1000	-2.283	-1.619	-1.264	1.339	1.717	2.457	0.037	1.016
2000	-2.251	-1.613	-1.254	1.327	1.706	2.451	0.035	1.009
5000	-2.296	-1.633	-1.281	1.291	1.658	2.356	0.013	1.000
10000	-2.315	-1.630	-1.274	1.299	1.671	2.369	0.016	1.005
ACR	-2.236	-1.645	-1.282	1.282	1.645	2.236	0.000	1.000

Note: Each set of finite sample critical values are based on $M = 50,000$.