New Scale-free Goodness-of-fit Tests for Rayleigh Distribution Based on Some Characterization and Some Special Property.

Ragozin Ilya

May 19, 2021

Characterization

Using one of the most famous characterization for exponential distribution family, which belongs to Desu M.M. (1971), we can get the characterization for Rayleigh distribution family:

Theorem

Let X and Y be positive independent identically distributed (iid) nonnegative random values(rv's) with a continuous df R. Then

$$X \stackrel{d}{=} \sqrt{2} \cdot \min(X, Y) \tag{1}$$

if and only if R belongs to the Rayleigh family of df's with scale parameter $\sigma > 0$ having the density $r(x, \sigma), \theta \in \mathbb{R}$, where

$$r(x, \sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}. \tag{2}$$

Building of statistics

We have built two criteria based on this characterization:

$$IU_{1,n} = \int_0^\infty (F_n(t) - U_{1,n}(t))dF_n(t)$$

and

$$KU_{1,n} = \sup_{t>0} |F_n(t) - U_{1,n}(t)|,$$

where $F_n(t)$ is the usual edf, and $U_{1,n}(t)$, $t \in \mathbb{R}$, is the U-empirical df for the $\sqrt{2} \cdot \min(X, Y)$:

$$U_{1,n}(t) = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \mathbb{1}\left\{\sqrt{2} \cdot \min(X_i, X_j) < t\right\}$$

Special property

Below we consider a special property of the Rayleigh family and build another two tests.

Theorem

Let X,Y be two nonnegative iid rv's with a df F, which belongs to the Rayleigh family. Then $Z = \frac{X}{Y}$ has the following df:

$$F_Z(t) = \frac{t^2}{1+t^2}, t \leq 0 \tag{3}$$

(Massom Ali, Nadarajah, Woo (2005))

In fact this property in not a characterization of Rayleigh law, because it is possibly to construct non-Rayleigh densities for X and Y under which X/Y still has the same df. Thereby, this will lead to inconsistency of tests based on this property against certain alternative. However, according to the usual concepts of testing statistical hypotheses, the evidence can be sufficient only for the rejection of the null-hypothesis H_0.

Building of statistics

Now construct two criteria based on this special property:

$$IU_{2,n} = \int\limits_{o}^{\infty} \left(U_{2,n}(t) - \frac{t^2}{1+t^2} \right) te^{-\frac{t^2}{2}} dt,$$

and

$$KU_{2,n} = \sup_{t>0} \left| U_{2,n}(t) - \frac{t^2}{1+t^2} \right|,$$

where$F_n(t)$ is still the usual edf, and $U_{2,n}(t)$, $t \in \mathbb{R}$, is the U-empirical df for X/Y:

$$U_{2,n}(t) = (c_n^2)^{-1} \left(\sum_{1 \leq i < j \leq n} \left(\frac{\mathbb{I}\{\frac{X_i}{X_j} < t\} + \mathbb{I}\{\frac{X_j}{X_i} < t\}}{2} \right) \right).$$

Building of statistics

Let $X_1, ..., X_n$ be i.i.d. observation with density f. We are testing null-hypothesis

$$H_0 : f(x) = r(x, \sigma), \sigma > 0,$$

where r(x,·) is the density for Rayleigh family (2). Previously we constructed four criteria, which will use for testing the null-hypothesis H_0 against some close alternatives described below. Also according to the Glivenko-Cantelli theorem for U-empirical df's, under H_0 the statistics IU_n and QU_n should be small.

Bahadur's Theorem

We will compare our tests using the well-known Bahadur efficiency. The crucial notion there is the so-called exact slope.

Theorem

Bahadur Theorem.
Suppose that the sequence of statistics T_n satisfies the following conditions:

1. $T_n \xrightarrow{\mathbb{P}_\theta} b(\theta)$, $\theta \in \Theta_1$, *where* $-\infty < b(\theta) < \infty$, *and*
2. $\lim\limits_{n\to\infty} n^{-1} \ln \mathbf{P}_\theta(T_n \geq z) = -k(z)$ *for each* $\theta \in \Theta_0$ *and any* z *from an open interval* I, *on which function* k *is continuous and* $\{b(\theta), \theta \in \Theta_1\} \subset I$.

Then for all $\theta \in \Theta_1$, *the exact slope* $c_T(\theta)$ *exists and can be calculated as*

$$c_T(\theta) = 2k(b(\theta)).$$

Bahadur efficiency

Now we introduce the Kullback-Leibler information for the composite null-hypothesis for any alternative density $f(x, \theta)$:

$$K(\theta) = \inf_{\sigma > 0} \int_{-\infty}^{\infty} \ln \frac{f(x, \theta)}{r(x, \sigma)} f(x, \theta) dx, \tag{4}$$

where $r(x, \sigma) = \frac{x}{\sigma^2} exp(-\frac{x^2}{2\sigma^2})$. The exact slopes always satisfy the inequality

$$c_T(\theta) \leq 2K(\theta),$$

so the local Bahadur efficiency of the sequence of statistics T_n is defined as

$$eff_T = \lim_{\theta \to 0} \frac{c_T(\theta)}{2K(\theta)}. \tag{5}$$

Alternative densities

We present the alternative densities $f_i(x,\theta), \quad x \in \mathbb{R}, i = 1, 2, 3$

1. Weibull alternative with the density:

$$f_1(x,\theta) = \frac{(1+\theta)}{2^\theta} x^{2\theta+1} exp\left(-\frac{x^{2(1+\theta)}}{2^{1+\theta}}\right).$$

2. Lehman or Verhulst alternative with the density

$$f_2(x,\theta) = (1+\theta) f(x) F^\theta(x) = (1+\theta) x e^{-\frac{x^2}{2}} \left(1 - e^{-\frac{x^2}{2}}\right)^\theta$$

3. Gamma alternative with the density:

$$f_3(x,\theta) = \frac{x^{\theta+1} e^{-\frac{x^2}{2}}}{2^{\frac{\theta}{2}} \Gamma\left(\frac{\theta}{2}+1\right)}$$

4. Rice alternative with the density

$$f_4(x,\theta) = x \cdot exp\left(-\frac{(x^2+\theta^2)}{2}\right) I_0(x \cdot \theta),$$

where $I_0(\cdot)$–is the modified Bessel function of the first kind with order zero.

Kullback-Liebler information

Now express the Kullback–Leibler distance $K(\theta)$ between alternative and the composite null-hypothesis H_0 (see Eq. 4). Moreover, the infimum in equation 4 is obtained for $\sigma^2 = \frac{1}{2}\int\limits_0^\infty x^2 f(x,\theta)dx$.

Lemma

For a given density $f(x,\theta)$, when $\theta \to 0$

$$2K(\theta) = \theta^2\left(I_f(0) - \left(\int\limits_0^\infty \left(\frac{x}{\sqrt{2}}\right)^2 f'_\theta(x,0)dx\right)^2\right) + o(\theta^2),$$

where $I_f(0)$ is the Fisher information at zero for the density $f(x,\theta)$, which is equal to

$$I_f(0) = \int\limits_{-\infty}^{\infty} \frac{|f'_\theta(x,0)|^2}{f(x,0)}dx.$$

Kullback-Liebler information

However, this asymptotic θ^2 is not enough for Rice alternative, because $f'_{4,\theta}(x,0) \equiv 0$. In case of this densities we get the follow: $K_4(\theta) = \frac{1}{128} \cdot \theta^8 + o(\theta^8)$. For other alternatives collect the Kullback-Liebler information in table below:

Table: The Kullback-Liebler information as $\theta \to 0$.

Alternatives	f_1	f_2	f_3
$2K_i(\theta)$	$\frac{\pi^2}{6} \cdot \theta^2$	$\left(\frac{\pi^2}{3} - \frac{\pi^4}{36}\right) \cdot \theta^2$	$\left(\frac{\pi^2}{6} - 1\right) \cdot \theta^2$

Integral statistic $IU_{1,n}$.

Consider the auxiliary function $g(x,y;z) = \frac{1}{2} - \mathbf{1}\left\{\sqrt{2}\cdot\min(x,y) < z\right\}$.
The integral statistic $IU_{1,n}$ is asymptotically equivalent to the U-statistic of degree 3 with the centered kernel

$$\Phi_1(x,y,z) = \frac{1}{3}\left(g(x,y;z) + g(y,z;x) + g(x,z;y)\right).$$

Lets calculate the projection of this kernel:

$$\Psi_1(t) = \mathbb{E}\left(\Phi_1(X,Y,Z)|Z=t\right) = -\frac{1}{18} + \frac{1}{3}e^{-\frac{t^2}{2}} - \frac{4}{9}e^{-\frac{3t^2}{2}}, \quad t>0$$

Now calculate the variance of this projection:

$$\Delta_1^2 = E\Psi_1^2(X) \approx 0.00291,$$

Integral statistic $IU_{1,n}$.

Thereby our kernel Φ_1 is non-degenerate, and by Hoeffding's theorem one has:

$$\sqrt{n}\cdot IU_{1,n} \xrightarrow{d} \mathcal{N}\left(0, 9\Delta_1^2\right)$$

Moreover kernel Φ_1 is cental and bounded, we can describe the logarithmic large deviations of $U-$statistics with such kernels:

Theorem

For any $t>0$,

$$\lim_{n\to\infty} n^{-1}\ln\mathbb{P}(IU_{1,n}>t)=h_1(t)$$

where h is some continuous function such that
$h_1(t)\sim-\frac{t^2}{18\Delta_1^2}$, *as* $t\to 0$.

Integral Statistic $IU_{1,n}$. Bahadur efficiency.

Now we are able to calculate the local Bahadur exact slope $c_{IU_1}(\theta)$ of the sequence of statistics $\{IU_{1,n}\}$:

$$c_{IU_1}(\theta) \sim \frac{b^2_{IU_1}(\theta)}{9\Delta^2}, \quad \text{as} \quad \theta \to 0, \text{where} \quad b_{IU_1}(\theta) = \mathbb{E}_\theta\left(\Phi_1(X, Y, Z)\right).$$

Notice, for alternative $f_i, i = \overline{1,3}$ the asymptotic of $c(\theta)$ is θ^2, but for Rice alternative f_5 the asymptotic is θ^8. Thereby now we can calculate the local Bahadur efficiency for our alternative distributions:

$$eff_{IU_1}(f_i) = \lim_{\theta\to 0} \frac{c_{i,IU_1}(\theta)}{2K_i(\theta)}, \quad i = \overline{1..4}$$

Integral statistics $IU_{2,n}$.

The integral statistic $IU_{2,n}$ is asymptotically equivalent to the U-statistic of degree 2 with the centered kernel:

$$\Phi_2(x,y) = \frac{e^{-\frac{x^2}{2y^2}} + e^{-\frac{y^2}{2x^2}}}{2} - \frac{\sqrt{e}Ei(-\frac{1}{2})}{2} - 1,$$

where $Ei(-\frac{1}{2})$ is exponential integral in point -1/2. Calculate the projection of this kernel:

$$\Psi_2(t) = \mathbb{E}\left(\Phi_2(X,Y)|Y=t\right) = \frac{1}{2}\left(\frac{t^2}{1+t^2} + tK_1(t)\right) - \left(\frac{\sqrt{e}Ei(-\frac{1}{2})}{2} + 1\right)$$

where $K_1(t)$–modified Bessel functions of the second kind. Now calculate the various of this projection:

$$\Delta_2^2 = \mathbb{E}\Psi_2^2(X) \approx 0.000314$$

Integral statistics $IU_{2,n}$.

Thereby our kernel Φ_2 is non-degenerate, and by Hoeffding's theorem one has:

$$\sqrt{n} \cdot IU_{2,n} \xrightarrow{d} \mathcal{N}\left(0, 4\Delta_1^2\right)$$

Moreover kernel Φ is cental and bounded, we can describe the logarithmic large deviations of $U-$statistics with such kernels:

Theorem

For any $t > 0$,

$$\lim_{n\to\infty} n^{-1} \ln \mathbb{P}(IU_{2,n} > t) = h_2(t)$$

where h is some continuous function such that
$h_2(t) \sim -\frac{t^2}{8\Delta_2^2}$, *as* $t \to 0$.

Integral Statistics $IU_{2,n}$. Bahadur efficiency.

The statements for integral statistics $IU_{2,n}$ is quite similar with $IU_{1,n}$. So write it quite briefly: The local Bahadur exact slope $c_{IU_2}(\theta)$ of the sequence of statistics $\{IU_{2,n}\}$:

$$c_{IU_2}(\theta) \sim \frac{b^2_{IU_2}(\theta)}{4\Delta^2}, \quad \text{as} \quad \theta \to 0, \text{where} \quad b_{IU_2}(\theta) = \mathbb{E}_\theta\left(\Phi_2(X,Y,Z)\right).$$

Local Bahadur efficiency:

$$eff_{IU_2}(f_i) = \lim_{\theta\to 0}\frac{c_{i,IU_2}(\theta)}{2K_i(\theta)}, \quad i = \overline{1,4}.$$

Local Bahadur efficiency for $IU_{1,n}$ and $IU_{2,n}$.

Collect the exact local Bahadur slope and the values of Local Bahadur efficiency for our statistics in table:

Table: Local Bahadur efficies for integral statistics.

	$IU_{1,n}$		$IU_{2,n}$	
Alternatives	$c(\theta)$	eff	$c(\theta)$	eff
f_1	$1.1466 \cdot \theta^2$	0.697	$1.3197 \cdot \theta^2$	0.802
f_2	$0.4714 \cdot \theta^2$	0.807	$0.0506 \cdot \theta^2$	0.087
f_3	$0.1274 \cdot \theta^2$	0.198	$0.014 \cdot \theta^2$	0.022
f_4	$0.0023 \cdot \theta^8$	0.149	$0.0045 \cdot \theta^8$	0.288

Kolmogorov type statistics $KU_{1,n}$ and $KU_{2,n}$.

We return to the Kolmogorov type statistic $KU_{1,n}$. Unfortunately, its limiting behavior is unknown, but it is still possible to calculate the logarithmic large deviations asymptotics under H_0. These statistics are the supremum by t of the family of absolute values for $U-$ statistics with the kernels: for $KU_{1,n}$:

$$\Phi_1(X,Y;t) = \frac{1}{2}\left(\mathbb{1}\{x<t\} + \mathbb{1}\{y<t\}\right) - \mathbb{1}\left\{\sqrt{2}\cdot\min(x,y) < t\right\},$$

and for $KU_{2,n}$:

$$\Phi_2(X,Y;t) = \frac{1}{2}\left(\mathbb{1}\left\{\frac{x}{y}<t\right\} + \mathbb{1}\left\{\frac{y}{x}<t\right\}\right) - \frac{t^2}{1+t^2}.$$

Kolmogorov type statistics $KU_{1,n}$ and $KU_{2,n}$.

Lets calculate the projection for each kernel and the variance for each projection:

$$\Psi_1(s;t) = \mathbb{E}\left(\Phi_1(X,Y;t)|Y=s\right) =$$
$$\frac{1}{2}\left(\mathbb{1}\{s<t\} - e^{-\frac{t^2}{2}} - 1\right) + \mathbb{1}\left\{s > \frac{t}{\sqrt{2}}\right\} e^{-\frac{t^2}{4}},$$

and

$$\Psi_2(s;t) = \mathbb{E}\left(\Phi_2(X,Y;t)|Y=s\right) = \frac{1}{2}\left(e^{-\frac{s^2}{2t^2}} + 1 - e^{-\frac{t^2 s^2}{2}}\right) - \frac{t^2}{1+t^2}.$$

Now variances $\Delta^2_{KU_i} = \sup\limits_{t>0} \Delta^2_{KU_i}(t) = \sup\limits_{t>0} \mathbb{E}\Psi_i^2(X;t), \quad i=1,2;$

$$i=1, \quad \Delta^2_{KU_1} = \sup_{t>0}\left(\frac{1}{4}e^{-t^2}\left(e^{\frac{t^2}{2}} - 1\right)\right) = \frac{1}{16}; \text{ in point } t = \sqrt{2ln(2)}.$$

$$i=2, \quad \Delta^2_{KU_2} = \sup_{t>0}\left(\frac{(t-1)^2 t^2 (t+1)^2 (t^4+3t^2+1)}{4(t^2+1)(t^2+2)(t^2-t+1)(t^2+t+1)(2t^2+1)}\right)$$
$$= 0.00954; \text{ in points } t = 0.445 \quad \& \quad t = 2.257$$

Kolmogorov type statistics $KU_{1,n}$ and $KU_{2,n}$.

Thereby our families of kernels are non-degenerate and also centred and bounded, we can apply, under H_0, the theorem on logarithmic large deviation asymptotics for $U-$ empirical Kolmogorov statistics. We obtain the following result:

Theorem

for any $z > 0$ $\lim_{n\to\infty} n^{-1} \ln \mathbb{P}\{KU_{i,n} > z\} = k_i(z)$, *where* k_i *is some continuous function, such that* $k_i(z) \sim -\frac{z^2}{8\Delta^2_{KU_i}}$ *as* $z \to 0$.

The local Bahadur exact slope of $KU_{1,n}$ and $KU_{2,n}$.

Using this theorem we get the following formula for the exact slope of Kolmogorov type statistic: We get the following formula for the exact slope of Kolmogorov type statistic

$$c_{KU_i}(\theta, f) \sim \frac{b^2_{KU_i}(\theta)}{4\Delta^2_{KU_i}}\theta \to 0, \quad i = 1, 2.$$

where

$$b_{KU}(\theta) = \sup_{t>0} |b(\theta; t)| = \sup_{t>0} |\mathbb{E}_\theta \left(\Phi_1(X, Y; t)\right)|$$

Similarly with case of integral statistics the local exact slope for Rice alternative (f_5) has the asymptotics $\theta^8, \quad \theta \to 0$; for other alternatives asymptotics is common $\theta^2, \quad \theta \to 0$.

Bahadur efficiency of $KU_{1,n}$ and $KU_{2,n}$.

Collect the exact local Bahadur slope and the values of Local Bahadur efficiency for our statistics in table:

Table: Local Bahadur efficies for Kolmogorov type statistics.

	$KU_{1,n}$		$KU_{2,n}$	
Alternatives	$c(\theta)$	eff	$c(\theta)$	eff
f_1	$0.2601 \cdot \theta^2$	0.158	$1.3134 \cdot \theta^2$	0.798
f_2	$0.1054 \cdot \theta^2$	0.181	$0.518 \cdot \theta^2$	0.886
f_3	$0.028 \cdot \theta^2$	0.043	$0.564 \cdot \theta^2$	0.821
f_4	$0.0085 \cdot \theta^8$	0.544	$0.0038 \cdot \theta^8$	0.243

Conclusion

Let us collect all efficiencies of our tests and compare them:

Table: Local Bahadur efficiencies for test statistics.

	Alternatives			
	f_1	f_2	f_3	f_4
IU_1	0.697	0.807	0.198	0.149
IU_2	0.802	0.087	0.022	0.288
KU_1	0.158	0.181	0.043	0.544
KU_2	0.798	0.886	0.821	0.243

We can see that almost in all cases the test statistics based on special property are more efficient than the tests, based on characterization. However in case of Rice alternative the test is the most efficient. Also we can see that Kolmogorov type test, based on special property has a high efficiency, that is unexpected result in test comparison by efficiency.