A SURVEY OF TESTS FOR EXPONENTIALITY WITH APPLICATION TO HISTORIC DATA

Yakov Nikitin

Saint-Petersburg University, Russia

Lausanne, October 29, 2018

A review of exponentiality tests

Testing exponentiality is very important in Probability and Statistics, especially in reliability theory, queueing theory and survival analysis. First tests of exponentiality:

- Greenwood (1946)
- Sherman (1950)
- Moran(1951)
- Epstein (1953)

Later appeared hundreds of papers and numerous surveys on the subject, even two handbooks on exponential distribution:

- Balakrishnan and Basu (1995),
- Nabendu, Chun and Crouse (2002).

A review of exponentiality tests

Let $X_1, \ldots, X_n$ be a sample with df F. We are interested in testing the composite hypothesis $H_0 : F \in \mathcal{E},$ where

$$\mathcal{E} = \{F : \ F(x) = 1 - e^{-\theta x}, \ \ \theta > 0, \, x \geq 0\}$$

is the class of exponential distributions with unknown scale parameter θ. The alternative H_1 consists in that $F \notin \mathcal{E}$.

Often the class of alternatives is supposed to be narrower (IFR, DFR, NBU, NBUE, and other classes arising in statistical reliability theory.)
Consider the *failure rate*

$$r(t) = \frac{f(t)}{1 - F(t)}, \ t \geq 0.$$

The exponential law is the only law with $r(t) = const$, while the IFR class consists of distributions with *increasing* r, DFR class consists of distributions with *decreasing* r, etc.

A review of exponentiality tests

Often the alternatives are particular parametric families like Gamma, Weibull, Makeham, LFR (linear failure rate), etc.
For example, the Weibull density has the form

$$f(x;\theta) = (1+\theta)x^{\theta}\exp(-x^{1+\theta}),\ x \geq 0, \theta \geq 0.$$

The Makeham density is given by

$$g(x;\theta) = (1+\theta(1-\exp(-x)))\exp\{-x+\theta[x-(1-e^{-x})]\}, \quad x \geq 0, \theta \geq 0.$$

For $\theta = 0$ we get the exponential density.

A review of exponentiality tests

Consider some groups of tests for exponentiality.

- **Tests using characterizations.**

Exponential law has more characterizations than any other law.

The exponential law is the only solution of the "loss-of-memory" equation

$$\overline{F}(x+y) = \overline{F}(x)\overline{F}(y),\ \forall x, y \geq 0,$$

where $\overline{F} = 1 - F$. This implies the test statistic

$$\sup_{x,y\geq 0} |1 - F_n(x+y) - (1 - F_n(x))(1 - F_n(y))|,$$

with critical large values: Angus (1982), Koul (1986), Nikitin (1995). Here F_n is the empirical df.

A review of exponentiality tests

Some tests use other characterizations based on:

- order statistics and spacings;
- entropy properties and Kullback-Leibler information;
- mean residual life function $l(t) = E(T - t/T > t)$ for $t \geq 0$;
- empirical Laplace transform;
- empirical characteristic function, etc.

A review of exponentiality tests

II Normalized $L-$statistics

$$T_n = \frac{1}{n\overline{X}} \sum_{i=1}^{n} w_{i,n} X_{(i)},$$

where $X_{(i)}$ are the order statistics, while $w_{i,n},\ i = 1, \ldots, n$, are appropriate weights.

Examples:

1) Gini statistic (1914)

$$G_n = \frac{\sum_{i,j=1}^{n} |X_i - X_j|}{2n(n-1)\overline{X}} - \frac{1}{2}$$

corresponds to $w_{i,n} = 2\frac{i-1}{n-1} - \frac{3}{2}$.

2) Jackson statistic (1967) corresponds to $w_{i,n} = \sum_{s=1}^{i} \frac{1}{n-s+1}$.

3) Fortiana-Grané statistic (2002) appears when

$$w_{j,n} = 1 - \ln n - (n-j)\ln(n-j) + (n-j+1)\ln(n-j+1), j = \overline{1,n}.$$

A review of exponentiality tests

III **Scale-free tests based on ratios.**

Denote $U_i := X_i/\overline{X},\ i = 1 \dots n.$ Consider statistics

$$V_n = \frac{1}{n} \sum_{i=1}^{n} r(U_i),$$

where r is some real function on R^+.

Particular cases are:

- Greenwood statistic (1946) for $r(x) = x^2$;
- Moran statistic (1951) for $r(x) = \ln x$;
- Cox-Oakes statistic (1984) for $r(x) = (1-x)\ln(1-x)$, etc.

A review of exponentiality tests

IV Kolmogorov-type statistics.

We mean the test statistics of the form

$$J_n = \sup_{x \geqslant 0} \left| \frac{1}{n} \sum_{i=1}^{n} r(U_i,\ x) \right|,$$

where $r(s, x)$ is some function on $[0, 1]^2$.

Examples are the tests of Lilliefors (1969), of Baringhaus-Henze (2000) and of Khmaladze (1979 , 2008).

Using special properties of exponential law

Let X_1 and X_2 be i.i.d. exponential random variables.

Then the ratio $\frac{X_1}{X_2}$ has $F_{2,2}$ distribution.

This property, however, does not characterize the exponential distribution as there are more distributions with the same property.
Nevertheless, one can construct the goodness of fit tests in the same way as it were a characterization. For example, consider the test statistics

$$I_n = \int_0^\infty (\frac{t}{1+t} - H_n(t))e^{-t}dt, \; D_n = \sup_t |\frac{t}{1+t} - H_n(t)|.$$

which were proposed by Nikitin and Volkova (2016).
(Here $H_n(t)$ is the U-empirical df corresponding to the set of ratios $\frac{X_i}{X_j}$.)

A review of exponentiality tests

- How to compare all these tests?
- How can we choose the best and most efficient test?
- How do these tests work for real data?

There are two approaches:

A. Power comparisons via simulation for small and moderate samples. This approach involves extended tables of powers for different levels and alternatives, the plots of power functions are given.

B. Calculation of (asymptotic) test efficiency in Pitman or Bahadur sense. The second type of efficiency is most adequate as asymptotic normality of test statistics is not required.

Efficiency of tests

The main idea in the calculation of test efficiency belongs to Pitman (1948).

Suppose we have two competitive sequences of test statistics $\{T_n\}$ and $\{V_n\}$ corresponding to the level $\alpha \in (0,1)$. Denote $N_T(\alpha,\beta,\theta)$ and $N_V(\alpha,\beta,\theta)$ the sample sizes necessary to achieve the power $\beta \in (0,1)$ under the alternative θ. Define the relative (finite-sample) efficiency as

$$eff_{TV}(\alpha,\beta,\theta) = \frac{N_V(\alpha,\beta,\theta)}{N_T(\alpha,\beta,\theta)}.$$

This quantity is too complicated to be found. Therefore one finds

- Pitman efficiency $= \lim_{\theta\to\theta_0} \; eff_{TV}(\alpha,\beta,\theta)$;
- Hodges-Lehmann efficiency $= \lim_{\beta\to 1} \; eff_{TV}(\alpha,\beta,\theta)$;
- Bahadur efficiency $= \lim_{\alpha\to 0} \; eff_{TV}(\alpha,\beta,\theta)$;

The calculation of efficiencies is a hard task requiring the theorems on asymptotic normality and large deviations. Most of efficiencies for tests of exponentiality were found in 1995 - 2017 by Henze, Klar, Litvinova, Meintanis, Nikitin, Tchirina,Miloševic, Obradovic.

Lilliefors test

We can use the Kolmogorov-type statistic:

$$D_n(\hat{\theta}) = \sup_{t \geq 0} |F_n(t) - (1 - \exp(-t/\hat{\theta}))|,$$

where $\hat{\theta}$ is some estimator of θ. Lilliefors (1969) proposed the statistic

$$D_n(\overline{X}) := Li_n = \sup_{x \geq 0} \left|1 - F_n(x) - e^{-x/\overline{X}}\right|$$

with critical large values of Li_n.

As the limiting distribution of Li_n is non-normal and unknown, Pitman efficiency is not applicable, and we should use Bahadur efficiency.

The problem of the evaluation of Bahadur efficiency for test statistics $D_n(\overline{X})$ was open for many years. It was solved recently by Nikitin and Tchirina (2007).

Khmaladze test and its properties

Consider the parametric empirical process $\xi_n^*(x) = \sqrt{n}(F_n(x) - F(x, \hat{\theta}_n))$. Unfortunately, the limiting behavior of this process is very complicated and almost useless.

The idea of Khmaladze (1979) was to extract the martingale part from $\xi_n^*(x)$. Define the "compensated" empirical process as follows:

$$W_n(s) = \sqrt{n}(F_n(x) - K(x, F_n)), \quad s = F(x, \hat{\theta}_n),$$

where $K(x, F_n)$ is the so-called compensator from martingale theory. Under H_0

$$W_n(s) \longrightarrow_d W(s),\ n \to \infty.$$

Khmaladze test is based on $\sup_{s \geq 0} |W_n(s)|$ which converges to $\sup_{s \geq 0} |W(s)|$.

Khmaladze test and its properties II

There is the exact expression

$$P(\sup_{s\geq 0}|W(s)| < z) = \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\exp\{-\frac{\pi^2(2n+1)^2}{8z^2}\}.$$

For the exponential law the compensator was computed by Khmaladze and takes the complicated form (with $\hat{\lambda} = 1/\bar{X}$) :

$$K(x,F_n) = \frac{\hat{\lambda}}{n}\sum_{i:X_i\leq x}(2X_i - \frac{\hat{\lambda}}{2}X_i^2) + \hat{\lambda}(2+\frac{\hat{\lambda}}{2}x)x(1-F_n(x)) - x\frac{\hat{\lambda}^2}{n}\sum_{i:X_i>x}X_i.$$

We have calculated the approximate Bahadur efficiency of this test. The exact efficiency is still unknown.

Numerical values of B-efficiency

Look at the numerical values of Bahadur efficiency for 6 statistics and 4 alternatives to exponentiality.

Table: Values of Bahadur efficiency.

Test	Makeham	LFR	Weibull	Gamma
$Moran$	0,694	0,388	0,943	1
$Greenw$	0,75	1	0,607	0,388
$Gini$	1	0,75	0,876	0,694
$Jack$	0,75	1	0,607	0,388
$Lilli$	0,538	0,356	0,607	0,503
$Khmal$	0,75	1	0,607	0,388

nascency of scientific chronology

The first scientific chronology is attributed to Josephus Justus Scaliger (1540 – 1609), italian, who lived in France, Switzerland and Netherlands. His work was developed by Dionisius Petavius (1583 - 1652).

Fig.1 J. Scaliger

Treatment of historical data

The great Isaac Newton (1642 -1727) has also worked in chronology of history for 40 years (!), see his book "The Chronology of Ancient Kingdoms Amended" (London: 1728). He found numerous "controversies, repugnancies and uncertaincies" and gave surprisingly accurate estimates and averages for historical events.

Fig.2 I.Newton

Inscription on the Monument of Newton in Trinity College, Cambridge

Newton: Who surpassed the race of men in understanding.

[Lucretius]

Distribution of lifelengths

Now let apply the exponentiality tests to historical data.

It is well-known that **the life length of human beings cannot be exponentially distributed.**

Humans are aging, hence human lifetimes have an increasing failure rate, or force of mortality. As people age they are more likely to die. In contrast, exponentially distributed variables have constant failure rates.

It seems that the situation with historical rulers, kings, governors, etc.would be similar: they would have stopped ruling or reigning as a consequence of accumulated controversies and tensions plus damages of economic, social, politic or personal nature like it happens with human lives.

Rules of Roman emperors

However, Khmaladze, Brownrigg and Haywood in the recent paper "Brittle power: On Roman Emperors and exponential lengths of rule", *Statistics and Probability Letters, 77(2007), 1248 -1257,* proclaimed that

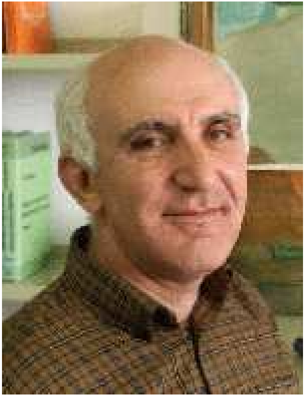


Fig. 3 Prof. E. Khmaladze

Lengths of rule of Roman Emperors

The lengths of rules of Roman Emperors are in agreement with exponential distribution.

Hence we are led to the surprising conclusion that their reigns ceased suddenly, "purely in random" in unexpected and unpredictable ways.

Reigns of Roman emperors

According to the chronology by Kienast (1990), the reigns of 76 Roman Emperors are as follows (in years):

Augustus	41,5
Tiberius	22,6
Caligula	3,8
Claudius	13,6
Nero	13,7
Galba	0,7
Otho	0,3
Vitellius	0,9
Vespasian	10,0
Titus	2,2
Domitian	15,0
Nerva	1,3

. . . .

Romulus Augustus 0,8

Fall of the Roman empire in 476 AD

Rules of Roman emperors

Using 53 lengths for the "decline period" from Nerva till Theodosius (395 AD) and applying the Khmaladze test, the authors concluded that the hypothesis of exponentiality is accepted.
They also found the same for :

- 161 Chinese emperors' from 221 BC to 1911 AD,
- 23 Spanish monarchs, starting with Isabella I in 1474,
- 24 Russian tsars, starting with Ivan IV the Terrible in 1547,
- 22 British kings, from Henry VII in 1483 to George VI (1936-1952).

Khmaladze et al. propose the following explanation of this.

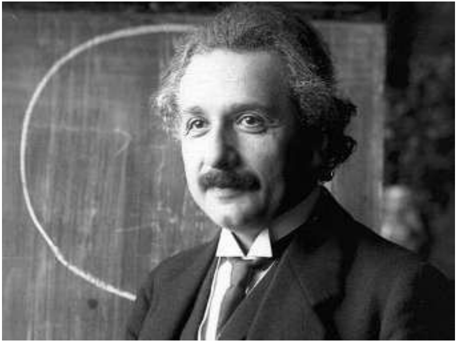
There are many challenges and threats for a reign, but there are also strong means to protect the ruler. Hence only exceptionally strong challenges could stop the reign. But in Probability, such occurrences form a Poisson process. Hence the inter-occurrences times would have an exponential distribution.

Do Roman emperors really have exponential reigns?

We have inspected the conclusions of Khmaladze et al. using numerous tests of exponentiality.
First consider the data of the decline period (96 - 395 AD) from Nerva till Theodosius they study and use exclusively the Khmaladze test. We found that some tests support the hypothesis of exponentiality at the level 0.05 while certain tests reject it.

Tests which accept the hypothesis:
Moran test, Greenwood test, Gini test, Epps-Pulley test,
Accept but scarcely:
Lilliefors test, Moran test
Reject:
Fisher test, chi-square test, Nikitin-Volkova test

Citation from Albert Einstein (1879-1955)



"No amount of experimentation can ever prove me right; a single experiment can prove me wrong."

(*Collected Papers of Albert Einstein, Vol. 7, Document 28. The Berlin Years: Writings, 1918-1921.*)

Rejection of the hypothesis

In case we test the hypothesis of exponentiality for all 76 Roman emperors, we also reject the null hypothesis by means of Gini test, Epps-Pulley test, Lilliefors test,Fisher test, chi-square test, only the Moran and Greenwood test accept this hypothesis.

Hence we disagree with Khmaladze, Haywood and Brownrigg. The reason is that possibly Khmaladze test is not sensitive and efficient enough to discover the deviation from exponentiality.

Another reason is that possibly the sample size is not large enough to apply the asymptotic formulas.

Further considerations

What happens with other dynasties of kings and emperors? It is an inexhaustible topic.

We have explored the reign lengths of Israel kings. There exists the famous hypothesis ascending to Morozov (Russian theoretician of terrorism who spent 25 years in the prison, book "CHRIST" in 7 vol. , 1924 - 1932)

Fig. 4 N.A.Morozov

Further considerations

and developed by M. Postnikov and A. Fomenko that in fact

Fig. 5 Prof. A. Fomenko

Israel kings = emperors of Western Roman empire,
Judean kings = emperors of Byzantinian empire !

Historic parallels

and the same sequence of rulers was counted TWICE(!!) by historians and chronologists.

Has the History created a duplicate of the first dynasty!?

Additional version:
Saul = Sulla, David = Julius Caesar, Solomon = Pompeius.

Observe that the division of Israel in two kingdoms: Judah and Israel after the death of king Solomon is in agreement with the irrevocable division of Roman Empire in two empires: the Western Roman Empire and the Byzantine Empire after the emperor Diocletian.

Astonishing parallel between Kings of Navarra and Sweden

See below the table of rules in Navarra (XI century) and Sweden (XVII century)

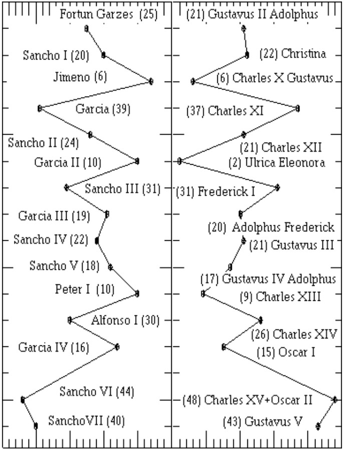


Fig.6 Kings of Navarra and Sweden

Kings of Israel and their rules

Kings of (undivided) Israel in order of reign:

Saul 24
David 42
Solomon 37

Kings of Israel in order of reign:

Jeroboam I 21
Nadab 1
Baasha 23
Elah 1,5
Zimri 0.02
Omri 11
Ahab 21
Ahaziah 1
Jehoram 11
Jehu 27
. . .
Hoshea 11

Assyrian captivity of Israel in 724 BC by Sargon II

Kings of Israel and their rules

Here the data is not very reliable as it corresponds to the events occurring 3000 years ago (king Saul ascended the throne in 1029 B.C.), and the only source of it is the Bible.

However, in all cases exponentiality is rejected, and the homogeneity of two samples (Western Roman emperors - Israel kings) is not supported.

The Life-Spans of Empires

Recently S.Arbesman (Harvard) studied the distributions of IMPERIAL lifetimes for more than three millenia. (Historical Methods, 44(2011), 127 - 129.)

Fig.7 Dr. S. Arbesman

The Life-Spans of Empires

Empire (41 species)	Adulthood(date)	Duration(centuries)
Elam(Mesopotamia)	1600 BC	10
Babylon (Mesopotamia)	1000 BC	2.5
Parthia (Iran)	60 BC	7.0
Rome (Europe)	0	4.5
Hun (Europe)	380	0.8
Yuen-Yuen (C.Asia)	400	0.3
Visigoth (Europe)	470	2.4

....

Arbesman found that their lifetimes undoubtedly follow the exponential law.

The Tyrants of Syracuse

Tyrants of Syracuse

Gelon I (491 BC-478 BC)
Hiero I (478 BC-466 BC)

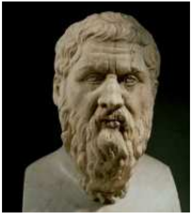


Fig.8 Dionysius II

....
Dionysius I, the Elder (405 BC-367 BC)
Dionysius II, the Younger (367 BC-357 BC)

....
Hippocrates and Epicydes (213 BC-212 BC)
The sample is surely non-exponential !

General judgment

The whole topic needs the detailed study and the use of additional tests of exponentiality.

The difference between History and Mathematics

When Kolmogorov (1903 -1987) was young (1921 -22),

Fig. 9 Acad. A.N.Kolmogorov

he was forced to make difficult choice between Mathematics and History which both strongly attracted him.

The difference between History and Mathematics

He attended the N.N. Luzin's seminar and constructed a Fourier series that diverges almost everywhere, gaining the praise of his tutor.

In the same time, he participated in the seminar of the Russian historian S.V. Bakhrushin, and there he reported his first research on the landholding in the Novgorod Republic in the 15-16 centuries. This research was highly estimated by the head of the seminar. However, he added that Kolmogorov has given only one proof of his conclusions, while History requires at least 5 (five!) different proofs.

The difference between History and Mathematics

The legend says that next morning Kolmogorov abandoned History FOREVER and committed himself to Mathematics.

Final statement

Thank you for your attention !