On the Coefficient of Variation and its Inverse

Giamloglou Georgia, Maleganou Athina, Papageorgiou Myrto, Farmakis Nikolaos
Department of Mathematics, Aristotle University of Thessaloniki

ABSTRACT

In this paper we are studying the coefficient of variation of a continuous random variable and some other concepts, like its inverse (symbol ICv or CV^{-1}) or its square inverse (symbol $ICV^2 = q$), e.tc. Basically, we try to develop the asymptotic sampling distribution of the inverse coefficient of variation $ICV = CV^{-1}$. This distribution is used to infer statistically significant results for the coefficient of variation or the inverse coefficient of variation of a random variable X without making a hypothesis for the population distribution of the variable X. We are focused on some cases of random variables following specific distributions dealing with the related parameters of those distributions. E.g. Our attention is mostly focused on extracting results (confidence intervals, hypothesis testing) for parameters of Gamma distribution, Weibull distribution e.tc.

Some examples are given in order to illustrate the particulars of the behavior of the ICv and ICV^2, related with the above-mentioned concepts: Confidence intervals, testing hypothesis, etc. for the random variable X and so on.

Key words: Inverse coefficient of variance, Sampling, Estimation, Hypothesis Testing

1. Introduction

The coefficient of variation (CV) is the standard deviation ratio to the mean value. CV has many applications that are used in statistical analysis. A fundamental property is that it helps analyzing the distribution of random variables, through sampling, with a very simple and easy to handle procedure. Distributions of random variables that are symmetrical, increasing and decreasing, as well as combinations of them have been studied, Farmakis (2003). The knowledge of the distribution of the CV and its inverse is calculated by sampling data, Farmakis (2003, 2010), Sharma & Krishna (1994). The distribution of the sampling CV is quite complicated and requires complex procedures to be determined. Thus, we often resort to its asymptotic form through a big sample, Sharma & Krishna (1994). Some symbols related to the distribution of CV and ICV are the following:

Symbols

μ, σ: mean, standard deviation for the population
$ICV = \frac{\mu}{\sigma}$, for the population inverse CV.
N : sample size
$\bar{x}, s$: sample mean, sample standard deviation
$c = \frac{\bar{x}}{s}$ the sample ICV
$\to^{L}$convergence in law
$\to^{P}$convergence in Probability
θ, β: parameters (scale, shape)
$ICV(b)$: inverse coefficient of variation versus parameter b
^: suggests that the parameter crowned by this mark is an estimator
$p(x)$: power function
$f(x)$: probability density function of X
$F(x)$: distribution function of the square X

2. Asymptotic distribution of ICv

We adopt the index $\hat{c}$ as a sample estimator of ICV, a symbol.
ICV sampling value properties:

- The method of the moments gives that $\hat{c}$ is an asymptotically unbiased and consistent estimator of ICV
- $\sqrt{N}(\hat{c} - ICV)$ has asymptotically standardized normal distribution

The second property is useful for testing hypotheses for CV or ICV.

Example:

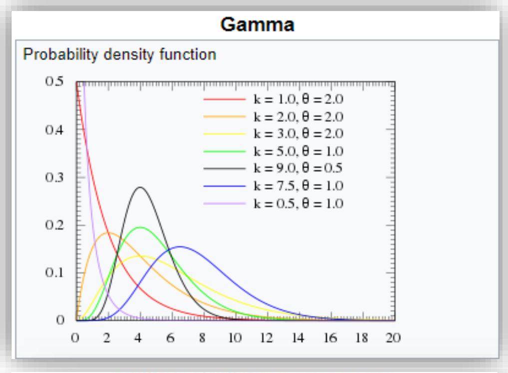
Apply to Survival / Renewal Theory (survival theory)
The widely used model is exponential because of the simplicity in its mathematical expression. Its major drawback, is that its constant failure rate, is rarely accurate.
For Gamma and Weibull:

$$\begin{cases} Gamma: \quad ICv = \sqrt{b}, \qquad Cdf\{x;\theta,b\} = gamf(\frac{x}{\theta};b) \\ Weibull: \quad ICv = \frac{G(1)}{[G(2)-G^2(1)]^{\frac{1}{2}}}, \; Cdf\{x;\theta,b\} = weif\left(\frac{x}{\theta};b\right), \\ G(m) = \Gamma(1+\frac{m}{b}) \end{cases}$$

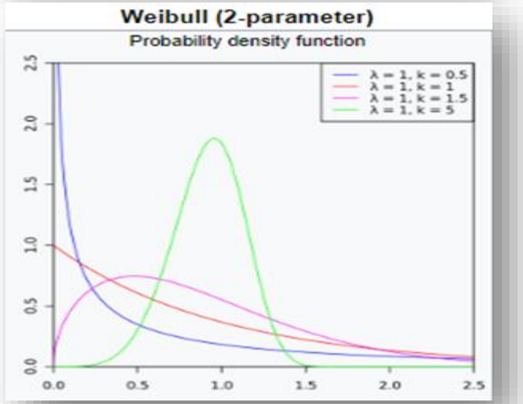
- Exponential distribution is a special case with $CFR = \frac{1}{\theta}$ for $b = 1$.
- ICV is 1-1 strictly increasing function of b.
- The problem of testing the hypothesis: $H_0: b = b_0$ is important and more difficult when θ is unknown. When $b_0 = 1$ the problem is limited to determining if the failure rate as a increasing, decreasing, or constant.

3. The case of Gamma and Weibull

Testing for b when θ is unknown requires the use of the statistic W. Although the distribution of W does not contain θ, it is quite complicated. In many reports, the distribution is discussed, and procedures are provided for the estimation of probabilities.



Another method for hypothesis testing of b contains $\frac{b}{\hat{b}}$. An empirical approach to the distribution of $\frac{b}{\hat{b}}$ is in the form: $g(\frac{b}{\hat{b}}) \sim c\sqrt{f(\cdot;h)}$, where g and h are selected so that mean and variance of both to agree.



5. Conclusions on b case

To avoid solving complicated equations of maximum likelihood, a large sample can be used for testing H_0: $b = b_0$ (with unknown θ) for Gamma distribution and Weibull distribution.

5.1 Point estimate

The function of inverse coefficient of variance is a 1-1 and ascending function of b. Therefore, testing $H_0: b = b_0$ is equivalent to the H_0: $ICV = ICV_0 = ICV\,(b_0)$ versus the alternative assumptions. Using the asymptotic sample distribution of inverse CV, which does not contain the variable θ, we can develop test over ICV using the Neyman-Pearson approach.

For control $H_0: b = b_0$, $H_1: b > b_0$ or $H_0: ICv(b) = ICv_0 = ICv(b_0)$
With rejection area: $c \geq ICv_0 + gauf^{-1}(a)$ with power function: $P(ICV) = gauf(z)$,
$z = \sqrt{N}\,(ICV_0 - ICV) + gauf^{-1}\,(a)$
The test is stable because $P \to 1$ when $N \to \infty$
Testing the equality of the parameters b (shape parameters) of the Gamma distribution or the Weibull distribution when the parameters θ (scale parameters) are unknown can be developed in a similar way.

5.2 Interval estimate

The 2-sided symmetrical $(1-a)$ confidence interval for ICv or b can be obtained as follows: $\Pr\{c - \varphi(a) \leq ICv \leq c + \varphi(a)\} = 1 - a$
$\varphi(a) = \frac{gauf^{-1}(\frac{1}{2}a)}{\sqrt{N}}$, as N increases so the width of the space decreases.

6. Exponentiality control

A special case of control is usually considered exponential ($b = 1$) versus IFR ($b > 1$) in Gamma and Weibull distributions.

7. Example with data

1.84870	.62880
.37610	.64610
.75000	1.65600
3.05300	1.71720
1.35450	1.93100
1.88020	1.05090
1.57000	1.61730
1.77080	1.31620
1.35920	.77050
3.04660	1.88890
1.79610	1.88890
1.53190	4.15050
.59030	

A random sample of 25 observations was created by Weibull distribution with $b = 2, \theta = 4$.

Remarks		
		Price
N	Validate Replies	25
	Absent Answers	0
Central and dispersion measures	Mean	1.6075880
	Variance	0.7290049
	Standard Deviation	0.85381783
	ICV	1.88287655
	25 percentage point	1.0509000
	50 percentage point	1.6173000
	75 percentage point	1.8802000
	Range	3.77440
	Minimum	.37610
	Maximum	4.15050

To test H_0: $b = 1$ versus the alternative H_1: $b > 1$ at a significance level $\alpha = 0.05$ with the rejection area of the inverse CV test: $(c-1)\sqrt{N} \geq 1.645$.
The sample observations show:

$N = 25, \bar{x} = 1.6076, s = 0.8538, c = 1.8829, (c-1)\sqrt{N} = 4.4145$

- $(c-1)\sqrt{N} = 4.4145 \geq 1.645$ thus, the inverse CV test rejects H_0 at a significance level of 0.05.
- The inverse CV test rejects H_0 at a significance level of 0.05 with $p = 0.76, b = 1.5$.

Using a 90% 2-sided symmetric confidence interval for ICV we get the interval [1.5928,2.2508] with a range of 0.658. The 90% 2-sided symmetric confidence interval for b is [1.635,2.2508] with a range of 0.760. Then, a 90% 2-sided symmetric confidence interval test for b is [1.474,2.520] with a range of 1,046. $ICV = c$ gives point estimate for $b = 2.118$ while the procedure $\hat{b} = 2.035$.

8. Highlights

- For Gamma distribution as well as for Weibull, reliability is a function of a, b. For known α, it is a 1-1 function of b and therefore a 1-1 function of $ICV(b)$. Therefore, $ICV's$ tests and confidence intervals can be used to develop tests and intervals for reliability.
- For lognormal distribution, where the standard deviation of $logx$ is b, conclusions can be drawn. $ICV = (exp(b^2) - 1)^{\left(-\frac{1}{2}\right)}$ that is independent of θ. So when θ is unknown, tests and confidence intervals can easily be extracted for the sample ICV.

BIBLIOGRAPHY

I. Φαρμάκης Ν. (2001) «Στατιστική, Περιληπτική Θεωρία-Ασκήσεις» Α & Π Χριστοδουλίδη, Θεσσαλονίκη.
II. Φαρμάκης Ν. (1915) «Δειγματοληψία& Εφαρμογές» Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα & Βοηθήματα, Κάλλιπος. Αθήνα.
III. Φαρμάκης Ν. (1916) «Εισαγωγή στη Δειγματοληψία» Αφοί Κυριακίδη Εκδόσεις Α.Ε. Θεσσαλονίκη.
IV. Farmakis N. (2003) "Estimation of Coefficient of Variation: Scaling of Symmetric Continuous Distributions", STATISTICs in TRANSITION, June 2003, Vol. 6, No 1,pp 83-96.
V. Farmakis N. (2010) "Coefficient of Variation: Connecting Sampling with some Increasing Distribution Models", Proceedings of Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA-20010) June 8-11, 2010 Chania Crete, Greece.
VI. Lawless J. F. (1982) "Statistical Models and Methods for Lifetime Data", John Wiley & Sons.
VII. Rao C. R. (1974) "Linear Statistical Inference and its Applications", John Wiley Eastern.
VIII. Sharma K. K. & Krishna H. (1994) "Asymptotic Sample Distribution of Inverse Coefficient-of-Variation and its Applications", IEEE Transactions on Reliability, Vol. 43, No 4, pp 630-633.