GAMMA-TYPE LIMIT DISTRIBUTION FOR EXTREME UNCENSORED ORDER STATISTICS UNDER RIGHT CENSORING

Dushatov N.T
Almalyk State Technical Institute
n_dushatov@rambler.ru

Abstract. We study the asymptotic distribution of statistics constructed from extreme order statistics under random right censoring. Classical results show that, for complete samples, suitably normalized gaps between the minimum and maximum order statistics converge to a Gamma distribution with shape parameter two. However, such results generally fail when censoring is present and observed extrema are distorted. In this paper, we propose a censored-data–adapted statistic based on the smallest and largest uncensored observations. Under mild regularity conditions and assuming independent censoring, we prove that the proposed statistic converges in distribution to a Gamma distribution with shape parameter two. The result provides a theoretically justified extension of classical extreme-value limit theorems to right-censored data and can be implemented using standard survival analysis tools.

Keywords. Right censoring; order statistics; extreme values; Gamma distribution; asymptotic distribution.

Annotatsiya. Ushbu maqolada tasodifiy o‘ng tomondan senzurlanish sharoitida ekstremal tartib statistiklari asosida qurilgan statistikalarining asimptotik taqsimoti o‘rganiladi. Klassik natijalarga ko‘ra, to‘liq tanlamalar uchun minimum va maksimum tartib statistiklari orasidagi mos ravishda normallashgan farqlar Gamma taqsimotiga (shakl parametri ikki) yaqinlashadi. Biroq senzurlanish mavjud bo‘lgan holatlarda kuzatilgan ekstremal qiymatlar buziladi va mazkur natijalar, odatda, o‘z kuchini yo‘qotadi.

Maqolada senzurlangan ma’lumotlar uchun moslashtirilgan, senzurlanmagan eng kichik va eng katta kuzatuvlarga asoslangan yangi statistika taklif etiladi. Mustaqil senzurlanish va yumshoq muntazamlik shartlari ostida ushbu statistikaning taqsimoti Gamma taqsimotiga (shakl parametri ikki) yaqinlashishi isbotlanadi. Olingan natija klassik ekstremal qiymatlar haqidagi limit teoremalarining o‘ng tomondan senzurlangan ma’lumotlar holatiga nazariy jihatdan asoslangan kengaytmasini beradi hamda omon qolish tahlilida qo‘llaniladigan standart statistik vositalar yordamida amalga oshirilishi mumkin.

Kalit so’zlar. O‘ng tomondan senzurlanish; tartiblangan statistiklari; ekstremal qiymatlar; Gamma taqsimoti; asimptotik taqsimot.

1. Introduction

Extreme order statistics play a fundamental role in probability theory and mathematical statistics, with applications ranging from reliability theory to survival analysis and risk assessment. Classical results show that, for independent and identically distributed observations from a continuous distribution, properly normalized statistics involving the minimum and maximum order statistics converge to exponential or Gamma-type limits [1,2].

In particular, it is well known that for a complete sample $X_1, X_2, \ldots, X_n$ with continuous distribution function F, the statistic

$$n\left(1+F\left(X_{(1)}\right)-F\left(X_{(n)}\right)\right)$$

converges in distribution to a Gamma distribution with shape parameter two as $n \to \infty$ [3]. This result relies critically on the joint asymptotic independence of the normalized lower and upper extremes.

In many applied problems, however, the underlying lifetimes or event times are not fully observable due to right censoring. In such settings, the maximum observed value is often censored, and classical extreme-value statistics based on observed data no longer reflect the behavior of the true extremes [4]. As a consequence, the aforementioned Gamma limit generally fails when naïvely applied to censored observations.

Several approaches have been proposed to adapt order-statistic–based methods to censored data, including inverse probability weighting and product-limit estimators [5,6]. Nevertheless, the asymptotic behavior of statistics involving extreme uncensored observations remains relatively less explored, especially in the context of Gamma-type limit distributions.

The aim of this paper is to fill this gap. We construct a statistic based on the smallest and largest uncensored observations and show that, after appropriate normalization, it converges to the same Gamma distribution as in the complete-data case. Our approach relies on a suitable transformation of the uncensored sample and exploits classical extreme-value arguments in a censored-data framework.

The remainder of the paper is organized as follows. In Section 2, we introduce the model and notation for right-censored data. Section 3 presents the main result, including the Gamma-type limit theorem for extreme uncensored order statistics. Proofs are provided in Section 4. Section 5 briefly discusses practical implementation and possible extensions.

2. Model and Preliminaries

We consider a standard right-censoring framework frequently used in survival analysis and reliability theory [4,6].

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed nonnegative random variables representing lifetimes or event times, with a continuous distribution function F and corresponding density f. Let $Y_1, Y_2, ..., Y_n$ be independent and identically distributed censoring times with distribution function G, independent of $\{X_i\}_{i=1}^n$.

The observed data consist of pairs $Z_i = \min(X_i, Y_i)$, $\delta_i = I(X_i \le Y_i)$, $i = \overline{1,n}$. Here, $\delta_i = 1$ indicates an uncensored observation, while $\delta_i = 0$ corresponds to a right-censored observation.

2.1 Distribution of Uncensored Observations

Let

$$m_n = \sum_{i=1}^{n} \delta_i$$

denote the number of uncensored observations. Under standard assumptions, $m_n \to \infty$ in probability $n \to \infty$ [8].

The conditional distribution of an uncensored observation satisfies

$$H(x) = \mathrm{P}(X \le x | \delta = 1) = \frac{\int_0^x f(t)\overline{G}(t)dt}{\int_0^\infty f(t)\overline{G}(t)dt}, \overline{G}(t) = 1 - G(t).$$

Thus, uncensored lifetimes do not follow the original distribution F, but instead follow a biased distribution H that depends on both the lifetime and censoring mechanisms [6].

Throughout the paper, we assume that H is continuous on its support.

2.2 Order Statistics of the Uncensored Sample

Let

$$X_1^u, X_2^u, ..., X_{m_n}^u$$

denote the uncensored observations, and let

$$X_{(1)}^u \le X_{(2)}^u \le ... \le X_{(m_n)}^u$$

be their order statistics.

Since H is continuous, the probability integral transform implies

$$U_i = H(X_i^u), i = 1, ..., m_n,$$

are independent and identically distributed random variables with

$$U_i \sim Uniform(0,1).$$

Consequently, classical results on order statistics for uniform samples can be applied to the transformed uncensored observations [1,2].

2.3 Extreme-Value Scaling

For a complete sample of size k from the uniform distribution, it is well known that

$$kU_{(1)} \Rightarrow Exp(1),\ k\left(1-U_{(k)}\right) \Rightarrow Exp(1),$$

and that the two limits are asymptotically independent as $k \to \infty$ [7].

In the censored-data setting, the effective sample size is random and equal to m_n. Under mild regularity conditions, including

$$\frac{m_n}{n} \xrightarrow{P} p,\ p = \mathrm{P}\left(X \le Y\right) > 0,$$

the asymptotic behavior of extreme uncensored order statistics is governed by m_n rather than n.

This observation motivates the construction of statistics based on the normalized extremes

$$m_n H\left(X_{(1)}^u\right),\ m_n\left(1-H\left(X_{(m_n)}^u\right)\right),$$

whose joint asymptotic behavior forms the basis of our main result.

3. Main Result

In this section, we establish a Gamma-type limit theorem for statistics constructed from the extreme *uncensored* order statistics under right censoring.

3.1 Assumptions

We impose the following standard conditions.

(C1) The lifetime variables $X_1, X_2, \ldots$ are i.i.d. with a continuous distribution function F.

(C2) The censoring variables $Y_1, Y_2, \ldots$ are i.i.d. with distribution function G, independent of $\{X_i\}$.

(C3) The probability of being uncensored is positive: $p = \mathrm{P}\left(X \le Y\right) > 0$.

(C4) The conditional distribution function

$$H\left(x\right) = \mathrm{P}\left(X \le x \middle| \delta = 1\right)$$

is continuous on its support.

Assumptions (C1)–(C4) are standard in survival analysis and ensure that the number of uncensored observations diverges as the sample size increases [4,8].

3.2 Definition of the Statistic

Let

$$X_{(1)}^u \le \ldots \le X_{(m_n)}^u$$

denote the order statistics of the uncensored observations, where $m_n = \sum_i^n \delta_i$.

We define the statistic

$$S_n = m_n\left(H\left(X^u_{(1)}\right) + 1 - H\left(X^u_{(m_n)}\right)\right).$$

This statistic measures the combined lower and upper tail mass of the uncensored sample under the transformed distribution H.

3.3 Theorem (Gamma-Type Limit under Right Censoring)

Theorem 3.1. Under assumptions (C1)–(C4), as $n \to \infty$

$$S_n \Rightarrow \Gamma(2,1),$$

that is, the statistic S_n converges in distribution to a Gamma random variable with shape parameter 2 and scale parameter 1.

Equivalently, for all $t \ge 0$,

$$\lim_{n\to\infty} \mathrm{P}\left(S_n \le t\right) = 1 - e^{-t}(1+t),$$

and the limiting density is

$$f(t) = te^{-t}, t \ge 0.$$

3.4 Discussion

Theorem 3.1 shows that the classical Gamma limit for extreme order statistics remains valid in the presence of right censoring, provided that the statistic is constructed from uncensored extremes and appropriately normalized by the random sample size m_n. It is important to emphasize that replacing $X^u_{(m_n)}$ with the largest observed value $Z_{(n)}$ generally destroys the Gamma limit, as the upper extreme is typically censored with non-negligible probability [4,6]. The proposed statistic avoids this difficulty by focusing exclusively on uncensored observations and by working under the induced distribution H. The result provides a rigorous justification for Gamma-type extreme-value approximations in censored-data settings and offers a natural bridge between classical order-statistic theory and modern survival analysis.

4. Proof of Theorem 3.1

This section provides a rigorous proof of the Gamma-type limit stated in Theorem 3.1.

4.1 Reduction to Uniform Order Statistics

Let

$$X^u_1, X^u_2, \ldots, X^u_{m_n}$$

denote the uncensored observations. By definition, their common distribution function is

$$H(x) = P(X \le x | \delta = 1).$$

Since H is continuous by assumption (C4), the probability integral transform yields

$$U_i = X^u_i,\ i = 1,2,\ldots,m_n,$$

Where

$$U_1,\dots,U_{m_n} \overset{i.i.d.}{\sim} Uniform(0,1).$$

Let

$$U_{(1)} \le \dots \le U_{(m_n)}$$

denote the order statistics of the transformed sample. By construction,

$$H\left(X^u_{(1)}\right)=U_{(1)}, \quad H\left(X^u_{(m_n)}\right)=U_{(m_n)}.$$

Therefore, the statistic

$$S_n = m_n\left(H\left(X^u_{(1)}\right)+1-H\left(X^u_{(m_n)}\right)\right)$$

can be rewritten as

$$S_n = m_n\left(U_{(1)}+1-U_{(m_n)}\right).$$

Thus, the problem reduces to studying the joint asymptotic behavior of the extreme order statistics of a uniform sample of random size m_n.

4.2 Asymptotic Behavior of Extreme Uniform Order Statistics

For a fixed sample size k, it is well known that

$$kU_{(1)} \Rightarrow Exp(1), \; k\left(1-U_{(k)}\right) \Rightarrow Exp(1), \; k\to\infty,$$

and that the two limiting variables are independent [1,7].

In our setting, the sample size m_n is random. However, by the law of large numbers,

$$\frac{m_n}{n} \xrightarrow{p} p > 0,$$

where $p = \mathrm{P}(X \le Y)$ by assumption (C3). In particular, $m_n \to \infty$ in probability as $n \to \infty$.

Standard arguments for triangular arrays imply that the classical extreme-value limits remain valid when the fixed sample size k is replaced by the random size m_n [8]. Consequently,

$$m_n U_{(1)} \Rightarrow Exp(1), \; m_n\left(1-U_{(m_n)}\right) \Rightarrow Exp(1).$$

4.3 Asymptotic Independence

To establish the joint convergence, we consider the joint density of $\left(U_{(1)}, U_{(m_n)}\right)$ conditional on $m_n = k$:

$$f_{U_{(1)},U_{(k)}}(u,v) = k(k-1)(v-u)^{k-2}, \; 0<u<v<1.$$

Define the scaled variables

$$A_n = m_n U_{(1)}, \; B_n = m_n\left(1-U_{(m_n)}\right).$$

Using the standard change of variables argument and letting $k \to \infty$, it follows that the joint density of (A_n, B_n) converges pointwise to

$$f_{A,B}(a,b) = e^{-a}e^{-b}, \; a > 0, b > 0,$$

which corresponds to two independent $Exp(1)$ random variables.

Hence,

$$(A_n, B_n) \Rightarrow (A, B),$$

where A and B are independent $Exp(1)$ variables.

4.4 Conclusion of the Proof

Since

$$S_n = A_n + B_n,$$

and the sum of two independent $Exp(1)$ random variables follows a Gamma distribution with shape parameter 2 and scale parameter 1, we conclude that

$$S_n \Rightarrow \Gamma(2,1).$$

This completes the proof of Theorem 3.1.

5. Conclusion

In this paper, we investigated the asymptotic behavior of statistics constructed from extreme order statistics under random right censoring. While classical Gamma-type limit results are well established for complete samples, their direct application fails in censored-data settings due to distortion of the upper extreme.

To address this issue, we proposed a statistic based on the smallest and largest uncensored observations and demonstrated that, after appropriate normalization by the number of uncensored observations, the statistic converges in distribution to a Gamma law with shape parameter two. The result holds under mild and standard assumptions, including independent censoring and continuity of the induced distribution of uncensored lifetimes.

Our findings provide a rigorous theoretical extension of classical extreme-value limit theorems to right-censored data and clarify the role of uncensored extremes in preserving Gamma-type asymptotic behavior. The proposed framework establishes a solid foundation for further methodological developments, including statistical estimation and simulation-based validation, which can be explored in future work.

References

1. David, H. A., Nagaraja H.N. Order Statistics. Wiley, 2003.

2. Arnold B.C., Balakrishnan N., Nagaraja H.N. A First Course in Order Statistics. SIAM, 2008.

3. Gnedenko B.V. The Theory of Probability. New York : Chelsea Pub. Co., 1968

4. Kleinbaum D.G., Klein M. Survival Analysis. Springer, 2012.

5. Kaplan E.L., Meier P. Nonparametric estimation from incomplete observations. *JASA*, 1958.

6. Fleming T.R., Harrington D.P. Counting Processes and Survival Analysis. Wiley, 1991.

7. Leadbetter, M. R., Lindgren, G., & Rootzén, H. Extremes and Related Properties of Random Sequences and Processes. Springer, 1983.

8. Van der Vaart, A. W. Asymptotic Statistics. Cambridge University Press, 1998.