Chapter 33
Order Statistics and Applications

E. Jack Chen
BASF Corporation, USA

ABSTRACT

Order statistics refer to the collection of sample observations sorted in ascending order and are among the most fundamental tools in non-parametric statistics and inference. Statistical inference established based on order statistics assumes nothing stronger than continuity of the cumulative distribution function of the population and is simple and broadly applicable. The authors discuss how order statistics are applied in statistical analysis (e.g., tests of independence, tests of goodness of fit, hypothesis tests of equivalence of means, ranking and selection, and quantile estimation). These order-statistics techniques are key components of many studies.

INTRODUCTION

Order statistics refer to the collection of sample observations sorted in ascending order and are among the most fundamental tools in non-parametric statistics and inference. Statistical inference established based on order statistics assumes nothing stronger than continuity of the cumulative distribution function of the population and is simple and broadly applicable. Important special cases of order statistics are the minimum and maximum value of samples, the sample median and other sample quantiles. Order statistics have been widely studied and applied to many real-world issues. It is often of interest to estimate the reliability of the component/system from the observed lifetime data. In many applications we want to estimate the probability that a future observation will exceed a given high level during some specified epoch. For example, a machine may break down if a certain temperature is reached. For more applications of order statistics, see (Arnold et al., 2008; David & Nagaraja, 2003; DasGupta, 2011).

It is the purpose of this chapter to present some of the more important results in the sampling theory of order statistics and of functions of order statistics and their applications, such as tests of independence, tests of goodness of fit, hypothesis tests of equivalence of means, ranking and selection, and quantile estimation. Order-statistics techniques continue to be key components of statistical analysis. For example,

quantile estimates are used to build empirical distributions, tests of goodness of fit are used to check the validity of the fitted input distribution, hypothesis tests of equivalence of means, and ranking and selection procedures are used to compare the performance of multiple systems.

BACKGROUND

Let $X_i, i = 1, 2, \ldots, n,$ denote a sequence of mutually independent random samples from a common distribution of the continuous type having a probability density function (pdf) f and a cumulative distribution function (cdf) F. Let $X_{[u]}$ be the uth smallest of these X_i such that $X_{[1]} \leq X_{[2]} \leq \ldots \leq X_{[n]}$. Then $X_{[u]}, u = 1, 2, \ldots n,$ is called the uth order statistic of the random sample $X_i, i = 1, 2, \ldots, n$. Note that even though the samples $X_1, X_2, \ldots, X_n$ are independently and identically distributed (iid), the order statistics $X_{[1]}, X_{[2]}, \ldots, X_{[n]}$ are not independent because of the order restriction. The difference $R = X_{[n]} - X_{[1]}$ is called the sample range. It is a measure of the dispersion in the sample and should reflect the dispersion in the population.

Suppose $U_1, U_2 \sim U(0, 1)$, where "$\sim$" denotes "is distributed as" and "$U(a, b)$" denotes a uniform distribution with range $[a, b]$. We are interested in the distribution of $U_{[2]} = \max(U_1, U_2)$, which can be viewed as the second order statistic. The cdf of $U_{[2]}$ is

$$
\begin{aligned}
F(x) &= P\Big(\max\big(U_1, U_2\big) \leq x\Big) \\
&= P\big(U_1 \leq x, U_2 \leq x\big) \\
&= P\big(U_1 \leq x\big) P\big(U_2 \leq x\big) \\
&= x^2
\end{aligned}
$$

Furthermore, the cdf of $U_{[1]} = \min(U_1, U_2)$, which can be viewed as the first order statistic, is

$$
\begin{aligned}
F(x) &= 1 - P\Big(\min\big(U_1, U_2\big) > x\Big) \\
&= 1 - P\big(U_1 > x\big) P\big(U_2 > x\big) \\
&= 1 - \big(1 - x\big)^2
\end{aligned}
$$

Consider a device contains a series of two identical components and the lifetime of the component $L_c \sim U(0, 1)$. If any of these two components fail, the device fails, such as chains. Then, the lifetime of the device $L_d \sim \min\big(L_{c1}, L_{c2}\big)$. On the other hand, if these two components are configured in parallel and the device fails only when both components fail, such as a strand contains a bundle of threads. Then the lifetime of the device $L_d \sim \max\big(L_{c1}, L_{c2}\big)$. These techniques can be used in reliability analysis to estimate the reliability of the component/system from the observed lifetime data.

Related Content

Effect of Personal Innovativeness, Attachment Motivation and Social Norms on the Acceptance of Camera Mobile Phones: An Empirical Study in an Arab Country

Kamel Rouibahand Hasan A. Abbas (2010). *International Journal of Handheld Computing Research (pp. 41-62).*

www.irma-international.org/article/effect-personal-innovativeness-attachment-motivation/48503

The M-Health Reference Model: An Organizing Framework for Conceptualizing Mobile Health Systems

Phillip Ollaand Joseph Tan (2009). *Mobile Computing: Concepts, Methodologies, Tools, and Applications (pp. 432-450).*

www.irma-international.org/chapter/health-reference-model/26519

Augmented Reality in Surgery: A New Approach to Enhance the Surgeon's Experience

José Inácio, João Ribeiro, Jaime Campos, Sara Silvaand Victor Alves (2018). *Next-Generation Mobile and Pervasive Healthcare Solutions (pp. 96-110).*

www.irma-international.org/chapter/augmented-reality-in-surgery/187518

Database Techniques for New Hardware

Xiongpai Qinand Yueguo Chen (2019). *Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics (pp. 546-562).*

www.irma-international.org/chapter/database-techniques-for-new-hardware/214642

Secure Deployment with Optimal Connectivity in Wireless Sensor Networks

Anju Sangwanand Rishipal Singh (2016). *International Journal of Mobile Computing and Multimedia Communications (pp. 1-21).*

www.irma-international.org/article/secure-deployment-with-optimal-connectivity-in-wireless-sensor-networks/161753