International Journal of Applied Mathematics & Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 5, Issue 6, Oct - Nov 2016; 37-42

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

ON FEW CONCEPTS OF RANDOM MEASUREMENTS

N. VADIRAJA

Assistant Professor, Statistics, Department of Community Medicine, Mysore Medical College and Research Institute Mysuru, Karnataka, India

ABSTRACT

Let $\{X_n, n\geq 1\}$ be a sequence of random variables. Many concepts such as sum of random variables, maximum and minimum of random observations and related statistics have been thought off and their properties have been studied in the literature. This paper gives few generalized concepts on random observations along with applications.

KEYWARDS: Max_ex, Moving Mex, Forward Moving Maxima, Backward Moving Minima

AMS Classification: Primary 60F15; Secondary 60G17

INTRODUCTION

The quality of highway depends on the area where it is constructed. Quality of highway constructed in the coastal area is very high as the temperature, weather and rain fall are extreme and hence can cause huge loss. The cost of highway in coastal area is very high compared to plain land. On the other hand same quality of highway in plain land is once again a loss to the government. As a result extreme behavior of temperature, pressure, weather and rain fall are very important factors to be considered in the decision of quality of highways. In the above example, let $\{X_n, n\geq 1\}$ be a sequence of random measurement with common distribution function (d.f) F. Let $F(x)<1$ for all x real. Gnedenko (1943) was first to study the degenerate limit of Y_n as $n \to \infty$, where $Y_n = \max(X_1, X_2, \ldots, X_n)$. The material strength is its weakest point This is an example of $M_n = \min(X_1, X_2, \ldots, X_n)$. Therefore it is meaningful to study the behavior of extreme values. Extensive discussion can be found in Galambos (1978).

This paper concentrates on defining the related concepts on random measurements with applications. Concepts defined are:

Forward moving maxima

Moving maxima for first k(n) observations

Moving minima for first k(n) observations

Forward moving minima

Max-ex for n observations

Moving max-ex for first k(n) observations

Backward moving max-ex

Forward moving max-ex

Moving mex for first k(n) observations

Backward moving mex

Forward moving mex

RESULT AND DISCUSSION

According to ICMR-GCP, subjects in a clinical trial can withdraw at any time. It is a right of the subject. As a result in clinical trials, drop out of subject is an issue. To overcome this issue, usually companies increase the estimated sample size by x% of drop outs to meet the required sample size. Suppose that at the entry point, say, n sample observations are selected and at the end point of clinical trial, k(n) subjects complete the trial. As a result k(n) is a fixed quantity. Hence, the clinical trial that has started with n subjects ends at k(n) subjects, i.e. it is a moving event. In social science research, it is common to discuss the response rate. Suppose for n subjects estimated, one may receive k(n) complete responses. As a result, n subjects moved to k(n) respondents. Hence due to drop out subjects, k(n) is a fixed quantity and is a real valued sequence varying from $2 \leq k(n) \leq n$. If there are no dropouts in clinical trials k(n) assumes value n. Clinical trial is meaningless if $k(n) \approx 0$. At the end of the study, after the dropouts, k(n) observations are available and this is a fixed constant. This fixed constant is a function of the sample size (n), where n is available in the beginning of the clinical trial. As a result, the assumption on k(n) naturally occurs. Any biological parameters such as biochemical parameters, pathological parameters and so on can be thought of for k(n) observations. In the case of hypertension and hyperglycemia it is customary to think of maximum value of the patient. As a result, Hebbar and Vadiraja (1996-97) have defined forward moving maxima as follows without application.

Forward moving maxima $Y_{k(n)} = \max(X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)})$ where k(n) is a sequence of positive integers, $2 \leq k(n) \leq n$ with certain assumption on k(n):

k(n) is non-decreasing

$\text{Sup}\,[k(n+1) - k(n)] \leq \mu$ (finite)

and

$k(n) = [n/(\log n)^{t(n)}]$ where $t(n) \rightarrow p$, $0 \leq p \leq \infty$ as $n \rightarrow \infty$

In the case of hypothermia, hypotension and hypoglycemia it is customary to think of minimum value of the corresponding parameter of the patient. In view of this the following concept is defined. Forward moving minima $Y_{k(n)}^{*} = \min(X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)})$. These k(n) random variables on to the right of n^{th} observation are different from those of backward moving minima and is defined as $V_{k(n)} = \min(X_{n-k(n)+1}, X_{n-k(n)+2}, \ldots, X_n)$ for k(n) observations to the left of n^{th} observations. The term moving maxima is due to Rothmann and Russo (1991).

Below few issues with k(n) observations in moving maxima will be addressed. Firstly, in moving maxima, the first k(n) observations are selected out of n observations. It appears that there is a biasedness among the selection of k(n) observations. Secondly, why can't k(n) observations be selected randomly? The answer is, as k(n) observations are selected from the randomly selected n observations from a population, the biasedness will not arise at all. The selection of k(n) observations randomly from randomly selected n observations is not necessary and leads to complications in dealing with the independence of events. As a result, the selection of random k(n) observations are not relevant. Also that, for larger n and n matching with N (the finite population size), k(n) observations also become larger. For larger $k(n) \leq n$, the concept of random selection loses its importance. Thirdly, what is the need for selection of k(n) observations from n? Fewer number of observations are studied in order to address constraints such as monetary, human resource, time and availability of resource issues. One can think of many applications those cannot be addressed over the entire population, especially in the case of research involving the destructive natured items. There are certain applications where in researcher will struggle hard to reach the required sample size also. For example in medical field, the replacement of artificial knee joint or replacement of bone with steel rod or organ transplantation and so on involve cost. In such situations one can think of k(n) observation concept or in the case of non funded research. Here the issue is that, does k(n) observations meet 80% power? As an answer, either one can think of n for more than 80% power and work k(n) for 80% or fix n for 80% power then this situation will be a particular case of k(n). Especially for larger n the k(n) can be thought off to achieve the better result. Also that, in market surveillance research in clinical trials of rejected drug, the company will be eager to know the factors for drug failure. In such situations they are not worried about n, and can concentrate on k(n). With all these explanation, k(n) concept is reasonable.

In a hospital based study, the researcher will be interested in the sequential sampling where-in observations on patients will be collected based on their occurrences. Now hyperthermia and hypothermia applications can be thought off. Define moving maxima of the first k(n) observations as $Y_{k(n)}^{**} = \max(X_1, X_2, \ldots, X_{k(n)})$ where X_i, i = 1,2,3,,,,k(n) be hyperthermia on i^{th} day or i^{th} individual. Observe that these random observations are different from that of forward moving maxima and backward moving maxima. On similar line, one can think of first k(n) observations on hypothermia application to study the minimum temperature of an individual on j^{th} day. In view of this, define moving minima of the first k(n) observations as $M_{k(n)}^{**} = \min(X_1, X_2, \ldots, X_{k(n)})$.

Next few concepts on the values not taken by the random phenomenon will be discussed. Let us consider a clinical trial application once again. Subjects/patients are very important to conduct the clinical trials. To select the required number of subjects it is customary to screen many volunteers. To enroll the volunteers to trial they have to pass the screening test. Let success be the clearing of screening test. The number of subjects required for the clinical trials is a fixed quantity. To meet with this number of success, doctors undergo screening tests on volunteers on several days. Suppose a subject doesn't pass the screening test, it is a miss hit. In other wards, failure is a miss hit by the success. On a given j^{th} day the number of miss hits by r successes is a random variable. It is customary by a pharmaceutical industry to invest on all volunteers to undergo screening test in selecting r number of success. Suppose a volunteer misses the success, the money invested on him will add to wastage of cost. Hence, company will be interested to know the maximum wastage cost. To address this question, it is meaningful to know the maximum of miss hits to success. As a result, following related concepts are defined. Over n number of days, the maximum excludent is defined as follows.

The maximum excludent (max_ex) over n observations:

$Z_n = \text{max\_ex}(X_1, X_2, \ldots, X_n)$

$= \max(t \geq 0: \text{for } X_j \neq t \text{ for } j = 1, 2, \ldots n)$.

Moving max_ex of first k(n) observations:

$Z_{k(n)}^{**} = \text{max\_ex}(X_1, X_2, \ldots, X_{k(n)})$

$= \min(t \geq 0: \text{for } X_j \neq t \text{ for } j = 1, 2, \ldots k(n))$.

Forward moving max_ex:

$Z_{k(n)} = \text{max\_ex}(X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)})$

$= \max(t \geq 0: \text{for } X_j \neq t \text{ for } j = n+1, n+2, \ldots n+k(n))$.

and

Backward moving max_ex:

$Z_{k(n)}^{*} = \text{max\_ex}(X_{n-k(n)+1}, X_{n-k(n)+2}, \ldots, X_n)$

$= \max(t \geq 0: \text{for } X_j \neq t \text{ for } j = n-k(n)+1, n-k(n)+2, \ldots n)$

To understand the max_ex concept an example is worked out below. Let X_1, X_2, X_3 are random variables representing "number of miss hit to success on j^{th} day", j=1,2,3. Let the number of success i.e. number of subjects cleared the screening test on each day be 3. Let the number of miss-hit to success on each day be 13, 30, 15. Then max_ex is:

$Z_n = \max(t \geq 0: \text{for } X_j \neq t \text{ for } j = 1, 2, 3)$

$= \max(X_1 \neq 3, X_2 \neq 3, X_3 \neq 3)$

$= \max(X_1 = 13, X_2 = 30, X_3 = 15)$

Note that the company will also be interested in estimating the minimum wastage cost. Hence minimum of the number of miss hit to success is also important. As a result, when we concentrate on the k(n) number of days, k(n) being a fixed quantity, for different k(n) the minimum of miss-hit random variables also changes. As a result forward moving minimal excludent (moving mex) is meaningful. The term mex is due to Conway(1978). In view of this, define forward moving mex as:

$L_{k(n)} = \text{mex}(X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)})$

$= \min(t \geq 0: \text{for } X_j \neq t \text{ for } j = n+1, n+2, \ldots n+k(n))$.

On similar lines backward moving mex is defined as:

$L_{k(n)}^{*} = \text{mex}((X_{n-k(n)+1}, X_{n-k(n)+2}, \ldots, X_n)$

$= \min(t \geq 0: \text{for } X_j \neq t \text{ for } j = n-k(n)+1, n-k(n)+2, \ldots n)$

It is called backward moving mex because the k(n) random observations are left to the n^{th} observations. These observations are different from that of forward mex. Next define moving mex of first k(n) observations as:

$L_{k(n)}^{**} = \text{mex}((X_1, X_2, \ldots, X_{k(n)})$

$= \min(t \geq 0: \text{for } X_j \neq t \text{ for } j = 1, 2, \ldots k(n))$.

To understand the mex concept an example is worked out below. Let X_1, X_2, X_3 are random variables.

$L_n = \min(t \geq 0: \text{for } X_j \neq t \text{ for } j = 1, 2, 3)$

$= \min(X_1 \neq 0, X_2 \neq 3, X_3 \neq 6)$

$= \min(\{1,2,3,..\}, \{0,1,2,4,\ldots\}, \{0,1,2,3,4,5,7,\ldots\})$

CONCLUSIONS

Motto of this paper is just to define the newer concepts and to provide application. However, these concepts can be applied in many areas to the user's advantage.

REFERENCES

1. Conway, J.H.(1976). On numbers and games, Academic Press, New York.
2. Galambos, J. (1978). The asymptotic theory of extreme order statistics. John Wiley and Sons, New York.
3. Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum dune serie aleatire. Ann. Math., 44, 423 – 453.
4. Hebbar, H.V. and N. Vadiraja (1996-97). Almost sure limit points of sequences of moving maxima under power normalization, JISPS, 3, 17-34.
5. Rothmann, M.D. and R.P. Russo (1991). Strong limiting bounds for a sequence of moving maxima. Statist. Probab. Lett. 11, 403-410.