A Higher-Fidelity Frugal Quantile Estimator

Anis Yazidi[1(✉)], Hugo Lewi Hammer[1], and B. John Oommen[2]

[1] Department of Computer Science, Oslo and Akershus University
College of Applied Sciences, Oslo, Norway
anis.yazidi@hiao.no
[2] School of Computer Science, Carleton University, Ottawa, Canada

Abstract. The estimation of the quantiles is pertinent when one is mining data streams. However, the complexity of quantile estimation is much higher than the corresponding estimation of the mean and variance, and this increased complexity is more relevant as the size of the data increases. Clearly, in the context of "infinite" data streams, a computational and space complexity that is linear in the size of the data is definitely not affordable. In order to alleviate the problem complexity, recently, a very limited number of studies have devised *incremental* quantile estimators [7,12]. Estimators within this class resort to updating the quantile estimates based on the most recent observation(s), and this yields updating schemes with a very small computational footprint – a constant-time (i.e., $O(1)$) complexity. In this article, we pursue this research direction and present an estimator that we refer to as a Higher-Fidelity Frugal [7] quantile estimator. Firstly, it guarantees a substantial advancement of the family of Frugal estimators introduced in [7]. The highlight of the present scheme is that it works in the discretized space, and it is thus a pioneering algorithm within the theory of discretized algorithms (The fact that discretized Learning Automata schemes are superior to their continuous counterparts has been clearly demonstrated in the literature. This is the first paper, to our knowledge, that proves the advantages of discretization within the domain of quantile estimation). Comprehensive simulation results show that our estimator outperforms the original Frugal algorithm in terms of accuracy.

Keywords: Quantile estimation · Stochastic Point Location · Discretized estimation

1 Introduction

Estimation is probably the most fundamental and central problem in many areas of engineering and computer science. The entire training phase of classification deals with estimation in one way or the other. While solutions to estimating

---

the mean (and central or non-central moments) of a distribution have been well established for centuries, we consider the problem of estimating the quantiles of a distribution with minimal time and space requirements.

Apart from the phenomenon of estimation, there are three rather distinct computational paradigms that have emerged within the general area of computational intelligence listed below:

1. The first of these involves the Stochastic Point Location SPL problem [8] where the Learning Mechanism (LM) attempts to learn a point on the "line" when all that it receives are signals from a random environment, i.e., whether *it* is to the "Left" or "Right" of the unknown point. This point that the LM attempts to learn may be, for example, a parameter of a control system.
2. The second of these involves the concept of discretization. Unlike learning in a continuous probability space, it has been shown that in the field of Learning Automata (LA), it is advantageous to discretize the probability space. Discretized LA are, generally speaking, both faster and more accurate than their corresponding continuous counterparts.
3. The third of these are the unique issues encountered when one seeks to estimate the quantiles of a distribution rather than the mean or central/non-central moments of a distribution in an *incremental* manner.

Conceptually, the fundamental contribution of this paper is to present a single solution that represents the confluence of these three distinct paradigms.

2   On Enhancing the Frugal Estimator

Since our contribution falls into the family of *Incremental* Quantile Estimators, we now present an overview of this class of estimators.

2.1   Incremental Quantile Estimators

An incremental estimator, by definition, resorts to the last observation(s) in order to update its estimate. The research on developing incremental quantile estimators is sparse. Probably, one of the outstanding early and unique examples of incremental quantile estimators is due to Tierney, proposed in 1983 [10], and which resorted to the theory of stochastic approximation. Applications of Tierney's algorithm to network monitoring can be found in [4]. The shortcoming of Tierney estimator [10] is that it requires the incremental constructions of local approximations of the distribution function in the neighborhood of the quantiles, and this increases the complexity of the algorithm. Our goal is to present an algorithm that does not involve any local approximations of the distribution function. Recently, a generalization of the Tierney's [10] algorithm was proposed by the authors of [5], where the authors proposed a batch update of the quantile, where the quantile is updated every $M \geq 1$ observations.

In the same context of incremental estimators, Ma, Muthukrishnan and Sandler [7] recently devised an innovative incremental quantile estimator[1] called the Frugal scheme, that follows randomized rules of updates. The first algorithm presented in the manuscript of Ma, Muthukrishnan and Sandler [7] is a Frugal approach for estimating the median. The procedure for estimating the median is simple but also "surprising": One increments the estimate of the median by a fixed amount Δ ($\Delta > 0$) whenever the observation from the data stream is larger than the median, and decrements the estimate of the median by Δ whenever the observation is smaller than the corresponding estimate. Nevertheless, the Frugal algorithm presented later in the same manuscript in order to tackle any quantile estimate (apart from the median), is not a generalization of the median case. In fact, according to the general update equations, if we are attempting to find the 50% quantile (median) of the data stream, we need to increment up randomly with 50% probability (for observations larger than the median estimate) and decrement down randomly with 50% probability (for observations smaller than the median estimate). Thus, intuitively, the Frugal [7] algorithm fails to generalize the median case as we observe that the randomization is unnecessary for estimating the median. Moreover, we can intuitively infer that the Frugal algorithm will suffer also from the "unnecessary" randomization for quantile estimates that fall in neighborhood of 50%.

In [12], Yazidi and Hammer devised a truly multiplicative incremental quantile estimation algorithm. The main difference between that and the current work is that the latter algorithm operates on a continuous space, while this present work is in a discretized space.

When it comes to memory efficient methods that require a small storage footprint, histogram based methods form an important class. Viewed from this perspective, a representative work is due to Schmeiser and Deutsch [9] who proposed the use of equidistant bins, where the boundaries are adjusted online. Arandjelovic et al. [1] used a different idea than equidistant bins by attempting to maintain bins in a manner that maximizes the entropy of the corresponding estimate of the historical data distribution, and where the bin boundaries were adjusted in an online manner.

In [6], Jain et al. resorted to five markers so as to track the quantile, where the markers corresponded to different quantiles and the min and max of the observations. Their concept was similar to the notion of histograms, where each marker had two measurements, its height and its position. By definition, each marker had some ideal position, and some adjustments were made so as to keep it in its ideal position by counting the number of samples that exceeded the marker. Thus, for example, if the marker corresponded to the 80% quantile, its ideal position would be around the point corresponding to the 80% of the data points below the marker. Subsequently, based on the positions of the markers, the quantiles were computed by modeling it such that the curve passing through

---

[1] With some insight, one sees that this elegant median estimation procedure is similar to the Boyer and Moore algorithm [2] for computing the majority item in a stream, using only a single pass.

three adjacent markers was parabolic, and by using piecewise parabolic prediction functions[2].

Finally, it is worth mentioning that an important research direction that has received little attention in the literature revolves around updating the quantile estimates under the assumption that portions of the data are deleted. Such an assumption is realistic in many real-life settings where data needs to be deleted due to the occurrence of errors, or because they are out-of-date and thus should be replaced. The deletion triggered a re-computation of the quantile [3], which is considered a complex operation. The case of deleted data is more challenging than the case of insertion of new data, because data insertion can be handled easily using either sequential or batch updates, while quantile update upon deletion requires more complex update operations.

2.2 The Higher-Fidelity Frugal Estimator

To motivate our work, we concur with Arandjelovic et al. [1] who remark that most quantile estimation algorithms are not single-pass algorithms and are, thus, not applicable for streaming data. On the other hand, the single pass algorithms are concerned with the exact computation of the quantile and thus require a storage space of the order of the size of the data which is clearly an unfeasible condition in the context of "Big Data" streams. Thus, the work on quantile estimation using more than one pass, or storage of the same order of the size of the observations seen so far, is not relevant in the context of this paper. We also affirm the need for storage-constrained and single-pass algorithms.

In this article, we extend the results from Frugal [7] and present a Higher-Fidelity Frugal (H-FF) scheme where the median can be seen as an instantiation of our algorithm and not as exceptional case that requires a different set of rules. In addition, our H-FF scheme is shown to be faster and more accurate than the original Frugal scheme [7]. For the rest of the paper, in order to avoid confusion, we will refer to the original Frugal algorithm due to Ma, Muthukrishnan and Sandler [7], as the Original Frugal (OF). As mentioned earlier, our H-FF algorithm is based on the theory of Stochastic Point Location [8], and although the latter theory has found applications within discretized binomial and multinomial estimation in [13], as we shall see, its application here is unique. In addition, one can observe that the binomial/multinomial discretized estimators proposed by Yazidi et al. in [11,14] and Frugal [7] are similar. In fact, if we use the same update equations as in [11,14] with the "binary" observation being whether the current estimate sample is larger than the current estimate, then, interestingly, we obtain the OF scheme [7]!

---

[2] Clearly, though, such an approach would not be able to handle the case of non-stationary quantile estimation as the positions of the markers would be affected by stale data points.

Let $Q_i = a + i.\frac{(b-a)}{N}$ and suppose that we are estimating the quantile in the interval[3] $[a, b]$. Note $Q_0 = a$ and $Q_N = b$. Let Δ be $\frac{(b-a)}{N}$. Further, we suppose that the estimate at each time instant $\widehat{Q}(n)$ takes values from the $N + 1$ possible values, i.e., $Q_i = a + i.\Delta$, where $0 \leq i \leq N$.

For the sake of completeness, we will give the update equations for the OF algorithm introduced in [7]. Please note that the equations are slightly modified so as to obtain estimates within $[a, b]$. In addition, the step size Δ has a general form and is not limited to unity as done in [7].

$$\widehat{Q}(n+1) \leftarrow Min(\widehat{Q}(n) + \Delta, b), \text{ If } \widehat{Q}(n) \leq x(n) \text{ and } rand() \leq q, \quad (1)$$
$$\widehat{Q}(n+1) \leftarrow Max(\widehat{Q}(n) - \Delta, a), \text{ If } \widehat{Q}(n) > x(n) \text{ and } rand() \leq 1 - q, \quad (2)$$
$$\widehat{Q}(n+1) \leftarrow \widehat{Q}(n), \text{ Otherwise}, \quad (3)$$

where $Max(.,.)$ and $Min(.,.)$ denote the max and min operator of two real numbers while $rand()$ is a random number generated in $[0, 1]$.

Our H-FF algorithm has two different update equations depending on whether the quantile we are estimating is larger or smaller than the median.

Update equation for $q \leq 0.5$:

$$\widehat{Q}(n+1) \leftarrow Min(\widehat{Q}(n) + \Delta, b), \text{ If } \widehat{Q}(n) \leq x(n) \text{ and } rand() \leq \frac{q}{1-q}, \quad (4)$$
$$\widehat{Q}(n+1) \leftarrow Max(\widehat{Q}(n) - \Delta, a), \text{ If } \widehat{Q}(n) > x(n), \quad (5)$$
$$\widehat{Q}(n+1) \leftarrow \widehat{Q}(n), \text{ Otherwise}. \quad (6)$$

Update equations for $q > 0.5$:

$$\widehat{Q}(n+1) \leftarrow Min(\widehat{Q}(n) + \Delta, b), \text{ If } \widehat{Q}(n) \leq x(n), \quad (7)$$
$$\widehat{Q}(n+1) \leftarrow Max(\widehat{Q}(n) - \Delta, a), \text{ If } \widehat{Q}(n) > x(n) \text{ and } rand() \leq \frac{1-q}{q}, \quad (8)$$
$$\widehat{Q}(n+1) \leftarrow \widehat{Q}(n), \text{ Otherwise}. \quad (9)$$

Theorem 1. *Let us assume that we are estimating the q-th quantile of the distribution, i.e., $Q^* = F_X^{-1}(q)$. Then, applying the updating rules given by Eqs. (4)–(6) for the case when $q \leq 0.5$, and Eqs. (7)–(9) when $q > 0.5$ yields: $\lim_{N \to \infty} \lim_{n \to \infty} E(\widehat{Q}(n)) = Q^*$.*

The proof of theorem is quite involved and is omitted here for the sake of brevity. The proof can be found in a unabridged version of this article [15].

---

[3] Throughout this paper, there is an implicit assumption that the true quantile lies in $[a, b]$. However, this is not a limitation of our scheme; the proof is valid for any bounded and probably non-bounded function.

2.3   Salient Differences Between the H-FF, SPL and OF

It is pertinent to mention that there are some fundamental differences between the H-FF and the SPL, both with regard to their *computational paradigms* and with regard to their respective *analyses*. There are also some fundamental differences between the H-FF and the OF schemes. We state them briefly below.

2.3.1   Differences Between the *Paradigms* of the H-FF and SPL

The following are the differences between the *paradigms* of the H-FF and SPL:

– Although the rationale for updating in the H-FF is *apparently* similar to that of the SPL algorithm [8], there are some fundamental differences. First, we emphasize that the SPL has a significant advantage. Indeed, the SPL assumes the existence of an "Oracle", the presence of which is, unarguably, a "bonus". In our case, since there is no "Oracle", the H-FF scheme has to simulate such an entity. Or more precisely, it has to infer the behavior of a fictitious "Oracle" from the incoming samples.
– Further, unlike the SPL, the H-FF has no specific LM either. The learning properties of the LM must now be encapsulated into the estimation procedure.

2.3.2   Differences Between the *Analyses* of the H-FF and SPL

The following are the differences between the *analyses* of the H-FF and SPL:

– From a cursory perspective, it could *appear* as if the Markov Chain that we have presented, and its analysis, are rather identical to those presented in [8]. However, although the similarities are few, the differences are more vital. The main differences are the following:
  1. First of all, unlike the original SPL, there is a non zero probability that in our present updating scheme, the estimate remains unchanged at the next time instant.
  2. As opposed to original SPL, in our case, the scheme never stays at the same state at the next time instant, except at the end states. Rather, the environment (our simulated "Oracle") directs the simulated LM to move to the right or to the left, or to stay at the same position.
– Unlike the work of [8], the probability that the "Oracle" suggests the move in the correct direction, is not constant over the states of the estimator's state space. This is quite a significant difference, since it renders our model to be characterized by a Markov Chain with state-*dependent* transition probabilities.
– A major advantage of this estimator and SPL-based estimators, in general, is that they are, by design, adequate to dynamic environments. In fact, the estimator is memory-less, and this is a consequence of the Markovian property. Thus, whenever a change takes place in the unknown underlying value of the target quantile to be tracked, our H-FF will *instantly* change its search direction since the properties of transition probabilities of the underlying random walk, change too.

2.3.3   Other Salient Differences Between the H-FF and OF

– Our H-FF is "semi-randomized" in the sense that only one direction of the updates is randomized and not both directions as in the case of the OF algorithm. In fact, whenever $q \leq 0.5$, we observe that the randomization is only applied for moving to the left (decrementing the estimate with probability $\frac{q}{1-q}$ which is less than unity). Similarly, when estimating a quantile q such that $q > 0.5$, the randomization is only applied for moving to the right (incrementing the estimate with probability $\frac{1-q}{q}$, which is again strictly less than unity).

– A fundamental observation is that for the median case, i.e., when $q = 05$, we obtain the Frugal update proposed as a exceptional case that deviates from the main scheme in [7] since $\frac{q}{1-q} = 1$. Formally, the median is estimated as follows:

$$\widehat{Q}(n+1) \leftarrow Min(\widehat{Q}(n) + \Delta, b) \text{ if } \widehat{Q}(n) \leq x(n), \tag{10}$$

$$\widehat{Q}(n+1) \leftarrow Max(\widehat{Q}(n) - \Delta, a) \text{ if } \widehat{Q}(n) > x(n). \tag{11}$$

3   Experimental Results

In order to demonstrate the strength of our scheme (denoted as H-FF), we have rigorously tested it and compared it to the OF estimator proposed in [7] for different distributions, under different resolution parameters, and in both dynamic and stationary environments. The results we have obtained are conclusive and demonstrate that the convergence of the algorithms conforms to the theoretical results, and proves the superiority of our design to the OF algorithm [7]. To do this, we have used data originating from different distributions, namely:

– Uniform in $[0, 1]$,
– Normal $N(0, 1)$,
– Exponential distribution with mean 1 and variance 1, and
– Chi-square distribution with mean 1 and variance 2.

In all the experiments, we chose a to be -8 and b to 8. Note that whenever the resolution was N, the estimate was moving with either an additive or subtractive step size equal to $\frac{b-a}{N}$. Thus, a larger value of the resolution parameter, N, implied a smaller step size, while a lower value of the resolution parameter, N, led to smaller step sizes. Initially, at time 0, the estimates were set to the value $Q_{\lfloor * \rfloor \frac{N}{2}}$. The reader should also note that an additional aim of the experiments was to demonstrate the H-FF's salient properties as a novel quantile estimator using only *finite* memory.

In this set of experiments, we examined various stationary environments. We used different resolutions, and as mentioned previously, we set $[a, b] = [-8, 8]$. In each case, we ran an ensemble of 1,000 experiments, each consisting of 500 iterations.

In Tables 1, 2, 3 and 4, we report the estimation error for the OF and H-FF for different values of the resolutions, N, for the Uniform, Normal, Exponential and Chi-squared distributions respectively. We catalogue the results for different values of the quantile being estimated, namely, q: 0.1, 0.3, 0.499, 0.7 and 0.9. From these tables we observe that the H-FF outperformed the OF in almost all the cases, i.e., for different distributions and for different resolutions. A general observation is that the error for both schemes diminished as we increased the resolution. For example, from Table 1, we see that the error for $q = 0.1$ decreased from 0.144 to 0.044 as the resolution increased from 50 to 500.

Table 1. The estimation error for the OF and H-FF algorithms for the Uniform distribution and for different values of the resolutions N and target quantiles.

q	0.1		0.3		0.499		0.7		0.9	
N	H-FF	OF	H-FF	OF	H-FF	F	H-FF	OF	H-FF	OF
50	0,144	0,144	0,197	0,198	0,245	0,246	0,220	0,220	0,176	0,175
100	0,104	0,103	0,146	0,146	0,160	0,161	0,157	0,159	0,122	0,122
150	0,074	0,075	0,121	0,122	0,135	0,137	0,128	0,131	0,100	0,101
200	0,069	0,068	0,106	0,107	0,117	0,120	0,113	0,115	0,088	0,089
250	0,063	0,063	0,096	0,097	0,106	0,109	0,102	0,106	0,081	0,083
300	0,055	0,056	0,089	0,090	0,098	0,104	0,096	0,102	0,080	0,082
350	0,051	0,052	0,083	0,085	0,091	0,097	0,094	0,099	0,081	0,084
400	0,050	0,050	0,078	0,081	0,088	0,095	0,091	0,098	0,082	0,086
450	0,046	0,047	0,075	0,077	0,083	0,091	0,089	0,098	0,083	0,087
500	0,044	0,044	0,072	0,075	0,082	0,091	0,088	0,097	0,084	0,089

A very intriguing characteristic of our estimator is that as the resolution increased, the estimation error diminished (asymptotically). In fact, the limited memory of the estimator did not permit us to achieve zero error, i.e., 100% accuracy. As noted in the theoretical results, the convergence centred around the smallest interval $[z\Delta, (z+1)\Delta]$ containing the true quantile. Informally speaking, a higher resolution increased the accuracy while a low resolution decreased the accuracy.

Another interesting remark is that both the OF and H-FF seemed to perform almost equally well for extreme quantiles, i.e., quantiles that are close to 0 or close to 1. However, as the true value of the quantile to be estimated became closer to 0.5, i.e., median, the H-FF had a markedly clearer superiority when compared to the OF.

The reader should note that the choice of 0.499 instead of 0.5 was deliberate in order to "avoid" using the exceptional rules presented with regard to the OF in [7], and that coincide with the rules of H-FF for the median. Thus, the estimation of the quantile for the value 0.499 was performed using the OF rules

Table 2. The estimation error for the OF and H-FF algorithms for the Normal distribution and for different values of the resolutions N and target quantiles.

q	0.1		0.3		0.499		0.7		0.9	
N	H-FF	OF	H-FF	OF	H-FF	F	H-FF	OF	H-FF	OF
50	0,341	0,339	0,376	0,377	0,361	0,358	0,377	0,376	0,956	0,956
100	0,259	0,259	0,258	0,260	0,251	0,250	0,259	0,258	1,030	1,042
150	0,235	0,239	0,210	0,213	0,205	0,203	0,212	0,212	1,082	1,096
200	0,229	0,236	0,188	0,192	0,176	0,175	0,190	0,191	1,122	1,133
250	0,233	0,244	0,171	0,175	0,157	0,156	0,170	0,175	1,154	1,170
300	0,242	0,258	0,161	0,165	0,144	0,142	0,160	0,168	1,187	1,204
350	0,254	0,272	0,152	0,162	0,133	0,129	0,152	0,159	1,216	1,237
400	0,273	0,293	0,148	0,155	0,124	0,120	0,148	0,158	1,245	1,273
450	0,290	0,310	0,143	0,155	0,116	0,113	0,144	0,154	1,277	1,302
500	0,305	0,329	0,142	0,154	0,112	0,109	0,142	0,152	1,303	1,332

Table 3. The estimation error for the OF and H-FF algorithms for the Exponential distribution and for different values of the resolutions N and target quantiles.

q	0.1		0.3		0.499		0.7		0.9	
N	H-FF	OF	H-FF	OF	H-FF	F	H-FF	OF	H-FF	OF
50	0,159	0,158	0,253	0,254	0,335	0,332	0,399	0,401	0,473	0,464
100	0,109	0,109	0,181	0,182	0,235	0,237	0,285	0,290	0,378	0,385
150	0,078	0,078	0,149	0,148	0,193	0,198	0,237	0,247	0,370	0,381
200	0,074	0,073	0,129	0,130	0,169	0,174	0,215	0,227	0,386	0,404
250	0,066	0,066	0,116	0,117	0,153	0,160	0,204	0,219	0,416	0,442
300	0,057	0,058	0,107	0,109	0,141	0,152	0,200	0,218	0,459	0,489
350	0,056	0,056	0,099	0,102	0,134	0,147	0,195	0,219	0,501	0,540
400	0,053	0,053	0,095	0,097	0,130	0,144	0,197	0,223	0,544	0,587
450	0,048	0,048	0,090	0,094	0,125	0,142	0,199	0,228	0,598	0,639
500	0,047	0,048	0,088	0,091	0,122	0,142	0,203	0,237	0,638	0,687

as per Eqs. (1)–(3) to avoid the unnecessary randomization of the OF around the median that could lead to higher errors, which was the earlier-mentioned shortcoming of the OF scheme.

Please note too that for the target values of the quantiles that were close to the initial point 0, the error was smaller than for those that are far away from initial point. Thus, for example, in Table 1, the error was lowest for the 10% quantile which is 0.1, which in this case, is closer to 0 than any other quantile in the the table, namely, 0.3 0.499, 0.7 and 0.9.

Table 4. The estimation error for the OF and H-FF algorithms for the Chi-squared distribution and for different values of the resolutions N and target quantiles.

q	0.1		0.3		0.499		0.7		0.9	
N	H-FF	OF	H-FF	OF	H-FF	F	H-FF	OF	H-FF	OF
50	0,088	0,088	0,254	0,254	0,348	0,345	0,453	0,454	0,600	0,606
100	0,063	0,063	0,149	0,149	0,234	0,231	0,322	0,326	0,519	0,525
150	0,051	0,052	0,126	0,125	0,192	0,192	0,270	0,272	0,535	0,567
200	0,045	0,045	0,105	0,104	0,167	0,170	0,245	0,253	0,597	0,638
250	0,040	0,040	0,094	0,095	0,150	0,153	0,227	0,243	0,686	0,731
300	0,037	0,036	0,085	0,085	0,139	0,142	0,220	0,238	0,765	0,822
350	0,033	0,033	0,079	0,079	0,129	0,136	0,218	0,239	0,842	0,915
400	0,031	0,031	0,074	0,075	0,122	0,128	0,220	0,244	0,933	0,987
450	0,029	0,029	0,070	0,070	0,118	0,125	0,218	0,254	1,003	1,062
500	0,027	0,027	0,067	0,068	0,113	0,121	0,222	0,258	1,073	1,134

4 Conclusion

This paper describes a scheme which is a confluence of three paradigms, namely, working with the foundations of Stochastic Point Location (SPL), the discretized world, and estimation of the quantiles in an incremental manner. We present a new quantile estimator which merges all these three concepts, and which we refer to as a Higher-Fidelity Frugal [7] (H-FF) quantile estimator. We have shown that the H-FF represents a substantial advancement of the family of Frugal estimators introduced in [7], and in particular to the so-called Original Frugal (OF) estimator.

Simulation results show that our estimator outperforms the OF algorithm in terms of accuracy.

References

1. Arandjelovic, O., Pham, D.S., Venkatesh, S.: Two maximum entropy-based algorithms for running quantile estimation in nonstationary data streams. IEEE Trans. Circuits Syst. Video Technol. **25**(9), 1469–1479 (2015)
2. Boyer, R.S., Moore, J.S.: MJRTY-a fast majority vote algorithm. In: Boyer, R.S. (ed.) Automated Reasoning: Essays in Honor of Woody Bledsoe, pp. 105–117. Springer, Netherlands (1991). doi:10.1007/978-94-011-3488-0_5
3. Cao, J., Li, L.E., Chen, A., Bu, T.: Incremental tracking of multiple quantiles for network monitoring in cellular networks. In: Proceedings of the 1st ACM Workshop on Mobile Internet Through Cellular Networks, pp. 7–12. ACM (2009)
4. Chambers, J.M., James, D.A., Lambert, D., Wiel, S.V.: Monitoring networked applications with incremental quantile estimation. Stat. Sci. **21**(4), 463–475 (2006)
5. Chen, F., Lambert, D., Pinheiro, J.C.: Incremental quantile estimation for massive tracking. In: Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 516–522. ACM (2000)

6. Jain, R., Chlamtac, I.: The P2 algorithm for dynamic calculation of quantiles and histograms without storing observations. Commun. ACM **28**(10), 1076–1085 (1985)
7. Ma, Q., Muthukrishnan, S., Sandler, M.: Frugal streaming for estimating quantiles. In: Brodnik, A., López-Ortiz, A., Raman, V., Viola, A. (eds.) Space-Efficient Data Structures, Streams, and Algorithms. LNCS, vol. 8066, pp. 77–96. Springer, Heidelberg (2013). doi:10.1007/978-3-642-40273-9_7
8. Oommen, B.J.: Stochastic searching on the line and its applications to parameter learning in nonlinear optimization. IEEE Trans. Syst. Man Cybern. Part B **27**(4), 733–739 (1997)
9. Schmeiser, B.W., Deutsch, S.J.: Quantile estimation from grouped data: the cell midpoint. Commun. Stat. Simul. Comput. **6**(3), 221–234 (1977)
10. Tierney, L.: A space-efficient recursive procedure for estimating a quantile of an unknown distribution. SIAM J. Sci. Stat. Comput. **4**(4), 706–711 (1983)
11. Yazidi, A., Granmo, O.-C., Oommen, B.J.: A stochastic search on the line-based solution to discretized estimation. In: Jiang, H., Ding, W., Ali, M., Wu, X. (eds.) IEA/AIE 2012. LNCS, vol. 7345, pp. 764–773. Springer, Heidelberg (2012). doi:10.1007/978-3-642-31087-4_77
12. Yazidi, A., Hammer, H.: Quantile estimation using the theory of stochastic learning. In: Proceedings of the 2015 Conference on Research in Adaptive and Convergent Systems, pp. 7–14. ACM (2015)
13. Yazidi, A., Oommen, B.J.: Novel discretized weak estimators based on the principles of the stochastic search on the line problem. IEEE Trans. Cybern. **46**(12), 2732–2744 (2016)
14. Yazidi, A., Oommen, B.J., Horn, G., Granmo, O.C.: Stochastic discretized learning-based weak estimation: a novel estimation method for non-stationary environments. Pattern Recognit. **60**(C), 430–443 (2016)
15. Yazidi, Anis Hammer L., H., Oommen, B.J.: Higher-fidelity frugal and accurate quantile estimation using a novel incremental (2017, to be submitted for publication). Journal version