The conduct of the sample average when the first moment is infinite

by J. L. MIJNHEER *

Summary Many books about probability and statistics only mention the weak and the strong law of large numbers for samples from distributions with finite expectation. However, these laws also hold for distributions with infinite expectation and then the sample average has to go to infinity with increasing sample size.
Being curious about the way in which this would happen, we simulated increasing samples (up to $n = 40000$) from three distributions with infinite expectation. The results were somewhat surprising at first sight, but understandable after some thought. Most statisticians, when asked, seem to expect a gradual increase of the average with the size of the sample. So did we. In general, however, this proves to be wrong and for different parent distributions different types of conduct appear from this experiment.
The samples from the "absolute Cauchy"-distribution are most interesting from a practical point of view: the average takes a high jump from time to time and decreases in between. In practice it might well happen, that the observations causing the jumps would be discarded as outlying observations.

1. Introduction

Let $\underline{x}_1, \ldots, \underline{x}_n$ be a sample from a random variable $\underline{x}$ with distribution function $F(x)$. Let further the expectation of $\underline{x}$ be infinite.
That means

$$\int_0^\infty x \, dF(x) = \infty \tag{1}$$

and

$$\int_{-\infty}^0 |x| \, dF(x) < \infty \tag{2}$$

For the sample average

$$\bar{\underline{x}}_n = \frac{\underline{x}_1 + \ldots + \underline{x}_n}{n} \tag{3}$$

follows

$$\bar{\underline{x}}_n \xrightarrow{\text{a.s.}} \infty \tag{4}$$

if the sample size increases. For a simple proof see [2].

Samples are simulated from the following three distributions:

a. absolute value of the Cauchy distribution.
b. Petersburg game.
c. return to origin.

a. Absolute value of the Cauchy distribution

The density is given by:

---

* Technical University, Twente.

$$f(x) = \begin{cases} \dfrac{2}{\pi}\dfrac{1}{1+x^2} & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases} \tag{5}$$

One can easily see that the expectation is infinite and the median is 1.

b. Petersburg game

A and B are tossing a true coin. A receives 2^r florins if it falls heads at the rth throw. If $\underline{x}$ is the gain of A at each trial, then the probability distribution of $\underline{x}$ is given by

$$P(\underline{x} = 2^r) = 2^{-r} \qquad r = 1, 2, \ldots \tag{6}$$

Consequently:

$$\mathscr{E}\,\underline{x} = \infty$$

and median of $\underline{x}$ is 2

c. Return to origin

A and B are tossing a true coin. At each throw A receives one florin if it falls heads and else he loses one florin. The probability of the number of throws between two successive returns to gain zero for A, is given by

$$P \text{ (numbers of throws} = 2m) = P_{2m} = \frac{1}{2m}\binom{2m-2}{m-1} 2^{-2m+2} \tag{7}$$

(see W. Feller [1]).

Application of Stirling's formula gives

$$P_{2m} \sim m^{-3/2}$$

So the expectation is infinite. The median of this distribution is 2.

2. Simulation

For simulating the samples from the three distributions we used pseudo-random numbers of the form $k/2^{26}$, with integer k and $0 \leq k < 2^{+26}$.

These numbers were generated by means of the function procedure RANDOM on the ELECTROLOGICA X-1 computer at the MATHEMATICAL CENTRE.

In the case of the absolute value of the Cauchy-distribution we need the property: If $\underline{x}$ is a random variable with a rectangular distribution on $[0,1]$ then

$$\underline{y} = \text{tg}\,\frac{\pi}{2}\,\underline{x}$$

has the probability distribution given by (5).

Proof

For $y \geq 0$ holds

$$P(\underline{y} \leq y) = P\left(\text{tg}\,\frac{\pi}{2}\,\underline{x} \leq y\right) = P\left(\underline{x} \leq \frac{2}{\pi}\,\text{arc tg}\,y\right) = \frac{2}{\pi}\,\text{arc tg}\,y$$

By differentiation follows

$$f(y) = \begin{cases} \dfrac{2}{\pi}\dfrac{1}{1+y^2} & \text{if } y \geq 0 \\[2mm] 0 & \text{else} \end{cases}$$

Thus, we simulate, in reality, a sample from a finite distribution, truncated at

$$a = \text{tg}\,\frac{\pi}{2}\,(1 - 2^{-26}) \approx 4.27 \times 10^7$$

This truncation is, however, negligible if the number of observations is 40000. For

$$P(\underline{y} > a) = 2^{-26}$$

and therefore the probability that one of the 40000 observations is greater than a is approximately 0.06%. The largest observation was in effect far from the point of truncation.

In both other cases we also used the procedure RANDOM to simulate the throws of a coin. In the figures the sample average is plotted against the number of observations. These figures were drawn by a plotter attached to the computer.

3. Discussion of the results

The first sample drawn is represented by the lower line in fig. 1 and reproduced on another scale in fig. 2. It is a sample from the absolute value of the Cauchy distribution and it does not really show a tendency to go to infinity at all. At the end of the experiment the average $\bar{x}_n$ seems to be nicely stabilized and in practice one would not suspect the parent population of having an infinite expectation. The second sample from the same population shows a big jump in the beginning, due to a very large
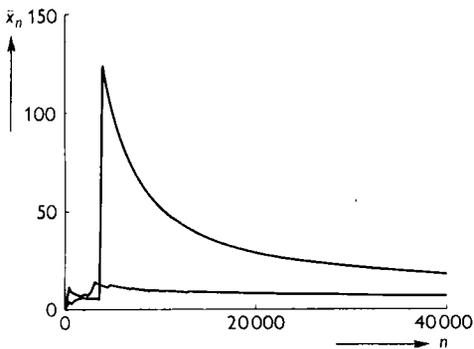

Fig. 1 Averages of two simulated samples from the absolute value of the Cauchy distribution plotted against the number of observations with step-size of 400 observations.
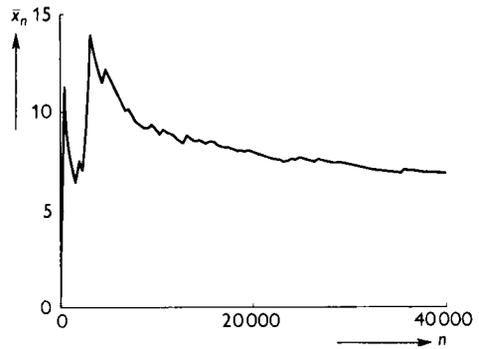
Fig. 2 Sample 1 of fig. 1 on another scale. Step-size of 400 observations.

observation, which would probably in practice be discarded as an outlying observation. But this is exactly the crux of the matter. Most of the observations from this distribution will be small (the median is 1!). Thus the sample „hesitates". It has to let its average go to infinity but, with so many small observations, that is difficult. Having reached a high value it will necessarily go down again. Thus it *has* to make high jumps from time to time to fulfill its destination but they are rare in this distribution. Nevertheless if the samples had been continued new jumps should have occurred. The second sample also seems to stabilize at $n = 40000$, but on another level. Thus the general pattern is an irregular sawtooth and the average is completely independable; but it is very much the question whether one would always realise this in practice.

The third sample, from the return to origin distribution shows another conduct, more in accord with the usually expected gradual increase of $\bar{x}_n$. The jumps due to large observations are much more common here. The difference between the two distributions also shows from the different order of magnitude of $\bar{x}_n$. For the return to origin $\bar{x}_n$ increases roughly linearly with n according to the formula

$$\lim_{n \to \infty} P\left(\frac{\bar{x}_n}{n} < x\right) = 2\left\{1 - \Phi\left(\frac{1}{\sqrt{x}}\right)\right\} \tag{11}$$

(see W. FELLER [1]), where Φ is the standard normal distribution function. For the absolute value of the Cauchy distribution one can show that

$$\frac{\pi}{2} \frac{\bar{x}_n}{\log n} \xrightarrow{P} 1, \tag{12}$$

thus the tendency to infinity of $\bar{x}_n$ is much weaker here.
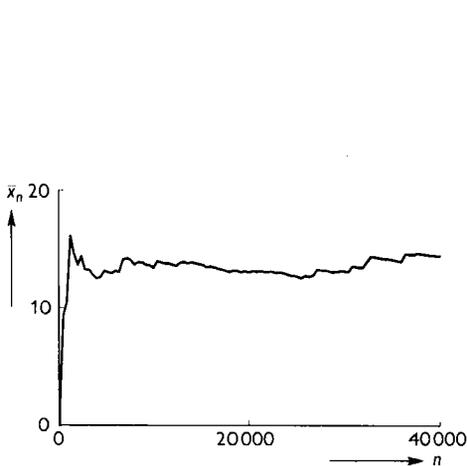

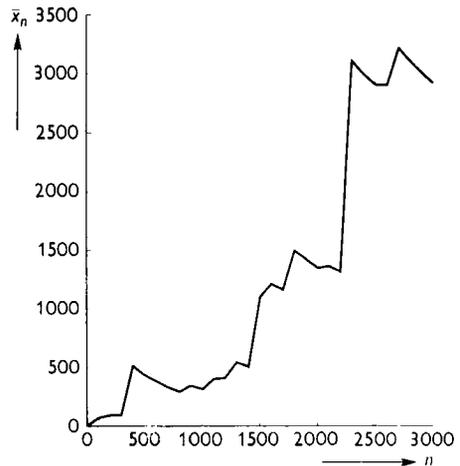Fig. 3 Sample average of the Petersburg game. Step-size of 400 observations.

Fig. 4 Sample average of the return to origin. Step-size of 100 observations.

(12) can be derived in the same way as, for the Petersburg game, the formula

$$\frac{\bar{x}_n}{2 \log n} \xrightarrow{P} 1 \tag{13}$$

is derived in FELLER [1]. From these three formulas the Petersburg game sample (fig. 3) would appear to be in an intermediage position between the other two. And indeed, the sample shows an intermediate conduct: $\bar{x}_n$ does not go down most of the time as in the Cauchy case, but it does not go up so nicely as in the case of the return to origin (fig. 4).

The question, whether the large observations in the absolute Cauchy case would perhaps in practice be discarded as outlying observations, will be considered in a later paper.

4. Acknowledgement

5. Literature

[1] FELLER, W., An introduction to probability theory and its applications, Vol. 1, John Wiley and Sons (1957).
[2] MIJNHEER, J. L., Oneindige verwachting en de wet van de grote aantallen, Statistica Neerlandica, Vol. 21 (1967), pp. 157–161.