A Matter of Means

Hopefully, this will put our discussion of Means in Hennesey and Patterson to bed so I can get some sleep. Now, the book and I are both write and wrong.

Types of Means

In statistics, that is, the derivation of characteristic values from populations or sets of values, there are many ways to compute an average or middle value. Consider a set of n values

$$X = \{x_1, x_2, x_3, \ldots, x_n\}.$$

The *arithmetic mean* of that set is given by

$$AM(X) = \frac{\sum_{i=1}^{n} x_i}{n}$$

The *harmonic mean* of set X is

$$HM(X) = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

The *geometric mean* of set X is given by

$$GM(X) = (\Pi_{i=1}^{n} x_i)^{\frac{1}{n}}$$

Exercise 1: Compute the following means for the values: $X = \{2, 1, 8, 5, 4, 2, 5, 1, 4\}$

1. $AM(X)$
2. $HM(X)$
3. $GM(X)$

Exercise 2: Suppose that we now define the set X_{inv} as the set produced by taking the reciprocal of the values in X. So, for X above, we have

$$X_{inv} = \{1/2, 1/1, 1/8, 1/5, 1/4, 1/2, 1/5, 1/1, 1/4\}$$

Now compute

1. $AM(X_{inv})$
2. $HM(X_{inv})$
3. $GM(X_{inv})$

Exercise 3: Suppose the values in set X represent task execution times. Then, what does the set X_{inv} represent? Based on the values computed in Ex. 1 and Ex. 2, what ordering relationship holds among the three means for each set of values.

Dealing with Time and Performance

The discussion of means in H&P is more imprecise than inaccurate, and thus more confusing than it should be. The *arithmetic mean of execution times* is introduced and defined correctly as a characterization of task execution time. The *harmonic mean of rates* is also defined correctly, from the formula standpoint. However, what confused me is the statement that the harmonic mean of rates can be used to track total execution time. In reality, what you are computing with the harmonic mean of rates is an average associated with throughput.

That is, the two means as discussed in H&P are tracking different characteristics of the target machine–average execution time and average throughput, respectively. The book is only incorrect in its implication, or rather the reader's inference that the two means were directly related. I apologize for any confusion this may have caused.

In conclusion, if the set X defined for Ex. 1 above represents task execution times, the resulting means refer to task execution time and are associated with response time. The set X_{inv} gives one set of task execution rates, and the means computed in Ex. 2 above are associated with performance, and track throughput.

Comparing Multiple Machine Performance

In H&P, the geometric mean is applied to the ratio of execution times on two machines. This is not meant to suggest that the geometric mean could not be used above to analyze performance or execution. Instead, the geometric mean is used for ratios, because the geometric mean of a ratio is equivalent to the ratio of the geometric means of the data points forming the numerator and denominator.

In contrast, the arithmetic mean cannot be used to analyze the ratios, not only because it doesn't "commute" with the division operation, but because it provides misleading results when analyzing ratios. The inability to use the arithmetic mean to compare ratios is the point of this subsection.

Overview

The discussion of means in H&P is there only to provide an explanation of different measures used to characterize task execution or performance in a general or average way for a given computer. The authors introduce this measure to point out its deficiencies when compared to the CPU time equation on which the ENTIRE BOOK is based. The point is, just because you can define and compute a number to characterize some facet of a computer architecture does not mean that the number is universally useful.