

Optimization of Coaching Using Six Sigma

Dumitru Colibaba-Evulet and Andrei Evulet

Department of Physical Education and Sports, University of Pitesti,
Str. Dragos; Voda nr 20, Pitesti 110186 Arges, Romania
Energy and Propulsion Technologies, GE Global Research,
ESB 243, Niskayuna, NY 12309, USA

Abstract: Problem statement: In today's sports world, coaching and training methods are the subject of intensive competition amongst rivals. Innovative technologies, methodologies and preparation methods are employed to achieve a sustainable competitive advantage. Effective performance measurement, statistical analysis and data based decision taking have been identified as key elements towards efficient performance management. The statistical approach has been emphasized in order to fill the gaps in athlete's preparation and improve performance. There is need to better connect the dots between the methods of statistical analysis and training and coaching methods used to improve the performance of an athlete or a team. **Approach:** Industry giants such as Motorola, General Electric and other improved the quality of their products through a statistically driven, six sigma methodology. Literatures in the six sigma performance measurement and improvement in industry, as well as traditional coaching in sports had been extensively reviewed. Research was carried out to establish potential application of six sigma in performance measurement and improvement in sports. **Results:** The six sigma approach was identified as a new method in measuring, analyzing, improving and analyzing sports performance and two examples were given. The methodology was applied on one example of data collected by authors over years of performance of a basketball team. Direction for further studies in the application of six sigma methodology in team and individual performance coaching and management were also identified. **Conclusion:** The results of this study determined that there is significant potential to successfully introduce the six sigma methodology of measurement, analysis and improvement in athletic performance. This new direction of analysis involved augmenting the existing traditional methods of coaching and performance management with a powerful methodology that clearly identifies root cause of athletic performance level, making it easier for the coach to take decisions and design a successful development plan.

Key words: Six sigma, quality scorecard, risk abatement, defects per million opportunities, closed-loop approach, statistical analysis, sports, basketball, athlete's performance, coaching process, olympic record, Upper Specification Limit (USL), Define, Measure, Analyze, Improve and Control (DMAIC), evaluation system

INTRODUCTION

"Six sigma", as applied by large corporations like Motorola, Honeywell and other, tries to create a closed-loop approach to minimization or elimination of all manufacturing defects generated by existing industrial processes (Pande *et al.*, 2000). In addition, the method can be equally applied to new processes through implementation in the design phase, thus avoiding bad decisions and errors committed in the past. Similarly, one could conceive that six sigma can be applied when preparing and coaching an athlete or a team of athletes before and during their competition season. Historical data gathered during the preceding competitions and

seasons become extremely important-the improvement of an athlete becoming a statistical problem to be solved by the coaching team, unbiased by the traditional approaches of physical training used in the past.

"Six sigma" is a methodology relying strongly on the preceding years' performance (Breyfogle, 2003). The data (performance, records and failures) recorded during the pre-season and the competitions are analyzed statistically. The analysis becomes unbiased by being transformed from a traditional coaching method to a statistical analysis of the failures, the importance of the factors involved and the down selection to the few vital elements that effect the athlete's or team's performance.

The essential actions to achieve the strategic goals of an athlete can be applied with precise, surgical accuracy using the statistical approach of this method.

MATERIALS AND METHODS

The materials used in this study are data collected from the performance records of a basketball team with sufficient statistical significance. The data was collected under the supervision of one of the authors while he was coaching the team. The statistical analyses performed on this data were used to draw conclusions of the development plan identified for the team and for the athletes (individual and team level). The statistical tools involved in this study were Minitab (Ryan *et al.*, 2005), analytical approaches (Breyfogle, 2003) and computational tools (Levine *et al.*, 2000; StatSoft, Inc., 2000). Because introduction of the six sigma methodology to sports needs to start with the building blocks for analyses, an a short description is listed here but further reading is advised in (Pande *et al.*, 2000, Breyfogle, 2003).

RESULTS AND DISCUSSION

The method has been used extensively in the last decades to improve the performance of final products by improving the quality of the study done by a business and understanding of the process variables that affect the final product (Pande *et al.*, 2000). In the end, the method inherently improves upon the productivity and eliminates the source of the defect generating process.

The assumption is that similarly to the business, the method can be equally applied to athletic activities with immediate benefits of better performance and reduction of performance variation (i.e., improving accuracy and improving repeatability by reducing variance) if identification and elimination of causes of failures or “defects”.

The goal of this study is to introduce, explain and extrapolate the six sigma methodology to the high performance sports research area by working on two examples.

The approach and application of six sigma (example 1: the 100 m flat runner): A 100 m flat runner may have mixed results during the season. Let us assume that the average time our athlete has per season is 10 sec, but that he targets to achieve or exceed constantly the Olympic record (Fig. 1). The coach has really 2 options:

- To prepare the athlete as before, based on his experience as a coach and the lesson learned in the past (the traditional approach)
- To use the six sigma method and determine, based on data collected in the past, the statistically proven right strategy and actions to be taken for achieving the goal by reducing the variance: This option can provide a consistent metric, set realistic goals and motivate the athlete

Figure 1 shows that the athlete needs to be able to “move” his average performance towards the goal (Olympic record time) while reducing his standard deviation (sigma) so that the size of the distribution (Fig. 2) narrows and results become more accurate and less spread. This is achieved only if “defective” or unacceptable performances, close to the USL (Upper Spec Limit-the worst performance recorded) are eliminated. This becomes a formula $Y = f(x)$, where Y is the performance factor-in this case the 100 m flat times-and the X’s are the inputs or variables thereof (like physical preparedness, mental status, wind, surface conditions). A pictorial approach is to use a “fishbone” graph to list the most important of these factors (cause and effect), showing graphically that the performances (Y) are a function of many factors X (X1.....Xn). The coach can add to the list of x’s the biologic side (X1), the psychological one (X2), the social interactions (X3), the technical preparedness (X4), the coaching process followed (X5, X6) with its elements and so on, all quantified. The performance reflected by the time to run 100 m flat is therefore a complex function of a large number of independent or quasi-independent variables (1-4....Xn), each contributing to the final variance of Y. All these variables need to be real numbers recorded from the past.

Six sigma is the process by which we push the performance average point towards the target (after determining, down selecting and eliminating the causes of failures to meet the target) and while reducing the variance (therefore increasing the accuracy of the results) until theoretically, if the athlete runs 1 million races in a season, his failures will total only 3.4 (meaning 3.4 times he fails to equal or break the Olympic record). This is in essence the definition of the six sigma approach.

The short form of a six sigma conversion table is shown in Fig. 5. Depending on the goals of the athlete, it shows the number of failures for any given sigma level, illustrated graphically to the right. These pictures can be used further when explaining and presenting the definitions of a few essential factors involved in applying this methodology:

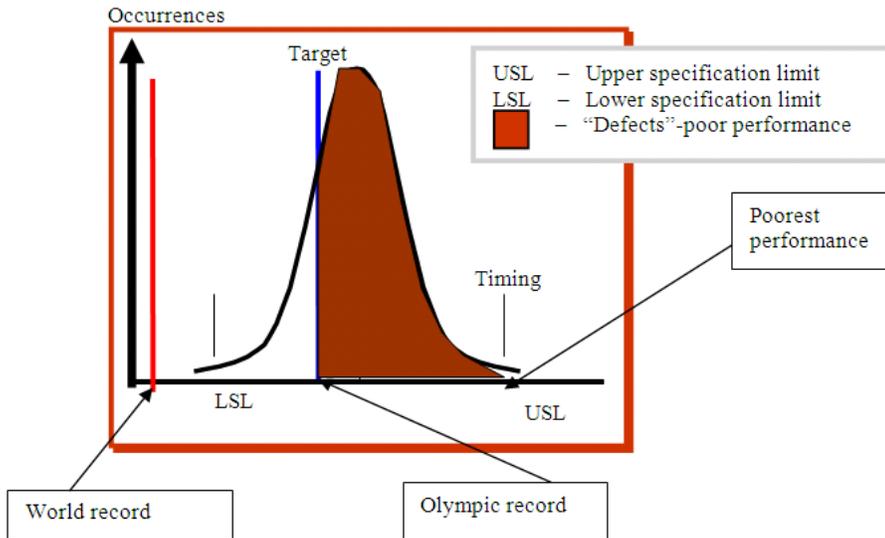


Fig. 1: 100 m flat variation example: current performance, before improvement

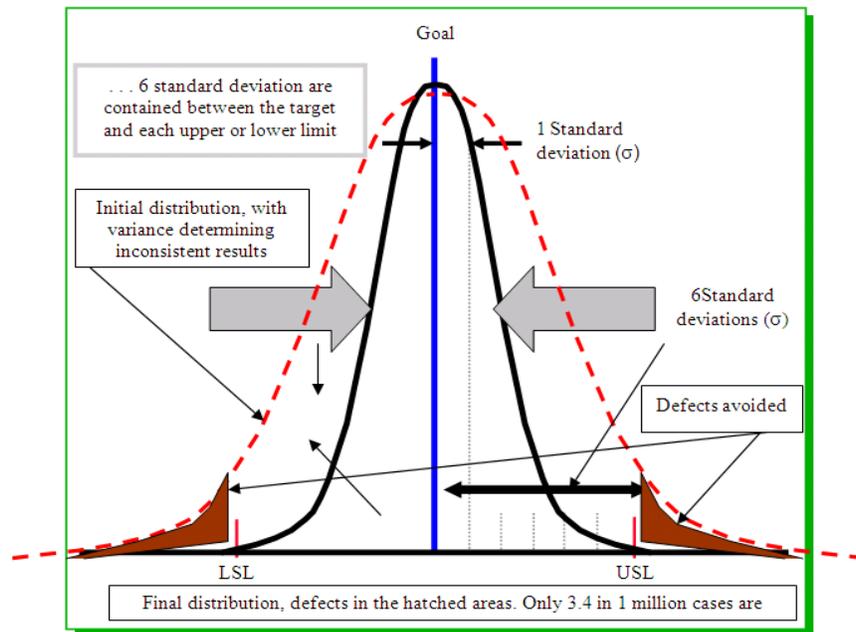


Fig. 2: Size of the distribution and six sigma illustration of the example

- DPMO = defects per million opportunities sauis the number of the failures an athlete had in meeting a target such as a personal record or an Olympic record, for instance, given one million opportunities; for example, 3 failures out of 10 attempts (or opportunities) means DPMO = 300.000 times in 1 million (Fig. 3) or roughly 2σ
- PC = Process Capability (Z) or how many times can the standard deviation fit between the accepted levels (LSL - USL), for the best distribution the athlete can produce/perform (Fig. 2)
- For Example: In order to improve from a 2 sigma level to a 4 sigma level, Z can be calculated for the upper and the lower spec limits as following $Z_{usl} = (usl - x_{med}) / \sigma$ and $Z_{lsl} = (x_{med} - lsl) / \sigma$, Fig. 3)
- USL = Upper Specification Limit is the upper acceptable limit, beyond which any result can be considered a "defect" (or failure to meet performance expected)

- LSL = Lower Specification Limit is similarly the lower acceptable limit-below which all results are considered defects (Fig. 2). It is important to mention that in the six sigma approach there may be only either one upper or one lower spec limit. For instance, the 100 m flat runner will have no LSL since any timing under the USL is considered a better performance. In this case no LSL exists, only an USL and a Z_{USL}
- LSL-USL= acceptable interval; defects are larger in value than the USL and lower than the LSL and are not acceptable. In sports, one of these limits is ignored because all timed sports (competing against time) have no lower spec limit while all sports where performance has other metrics (length, height) have no upper limits.
- $P(d)$ = probability of producing a defect-can be calculated as $P(d)$ -USL or as $P(d)$ -LSL = and is the sigma level of a process. The area under the curve to the right of Z is equal to the probability that a randomly selected member of the population will have a value equal to or greater than Z

The following probabilities are widely used as well:

$P(d)$ = Number of defect free opportunities/total opportunities, or

$P(d)$ = Number of defect opportunities/total opportunities

Defect = Results under or above the targeted goal, failure to meet expectations (e.g., running a 100 m race in more than 10 s). It is important to mention that the defect is independent of the cause or causes (may be multiple)

Data recorded through measurement and statistical evaluation can be of two types:

- Quantitative: Weight, time, length
- Qualitative: Example failed/pass or bad/good/outstanding-these data can be also given numerical values to better apply statistical tools

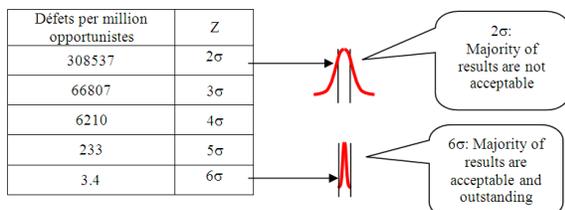


Fig. 3: Defect probabilities related to the sigma levels

Another important term is the defects per unit (dpu), which is the reverse of looking at DPMO.

The six sigma process comes in stages which need to be applied with rigorous discipline. The stages are: Define Measure, Analyze, Improve and Control (DMAIC), (Breyfogle, 2003, Olugu and Wong, 2009). Each of these steps has well established goals and contents as following:

- Define the process means establishing the realistic goal (Y), the present defects, the evaluation system, the rationale of the approach. During this phase the objectives are determined as well as the desired performance
- Measure: Record all the data; establish limits, statistical interpretation of the data, the control diagram and the action plan. The historical data and the current performance of the athlete are recorded during this phase
- Analysis-this phase is concerned with the understanding of the influence of the X factors upon the performance and with the hypothesis theory and testing, as well as the ranking of the factors of influence. This phase of analysis is a statistical approach, based on numbers, not on “gut feelings”. The analysis is performed with statistical tools (statistical tests, hypothesis evaluation) based on recorded data, unbiased-it is a mathematical root cause analysis which eliminates the traditional bias of the current approaches. Any loss of the objectivity during the analysis phase will render the method useless. Typically, a “Pareto” graph or a similar outcome show the few important factors that contribute to poor performance of an athlete
- Improve: In this phase we can eliminate the causes or improve the factors that affect performance on short or long terms and we can establish the entitlement to the best level the athlete can reach after improvement If the data analysis means identification and down selection of the vital few factors that cause failures, the improvement phase directs the coaching methods with maximum efficiency to the root causes and guarantees improvement of the athlete through efficient training. The process needs to be clearly understood and accepted by both the athlete and the coach
- Control-Improvement alone doesn't suffice: One needs to maintain control over any factors that may cause deviation from the improved performance. The progress made must be supervised and the efficiency of the method checked continuously, including major changes in the training process to correct unwanted deviations.

	1,2	3,4	Playoff
Regular season	1,2	3,4	Playoff
1. Offensive rebounds	2.4	2.3	3.6
2. Defensive rebounds	4	3.9	3.3
3. Intercepts	2.1	0.8	2.3
4. Steals	0.1	0.1	0
5. Blocks	0.2	0.3	0.3
6. 1x1	0.1	0.1	0
7. Winning balls	9.2	7.6	9.6
8. Wrong pass	1.4	0.9	1
9. Travel	0.2	0.2	0
10. Lost ball	1.3	0.9	1.6
11. 3 second rule	0.1	0.1	0
12. Lost balls	3	2.2	2.6
13. Assists	0.6	1.4	3
Free throws	78.00%	84.00%	88.00%
Long distance throws	36.00%	45.00%	81.00%
Semi throws	37.00%	48.00%	43.00%
In the paint throws	60.00%	61.00%	50.00%
Fast break	70.00%	66.00%	61.00%
Total offensive throws	0.48	0.55	0.53
14. Points made per game	15	14.8	21.6
Time played	28.3	27	33
Effective index	1.25	0.84	1.84
Total opportunities	39.7	35.6	48.9
Total lost opportunities	6	4.3	5.2
Index efficiency 1	84.89%	87.92%	89.37%
Index throws	50.75%	55.00%	58.75%
Index efficiency 2	71.21%	75.64%	78.71%

Annotations in the diagram:
 - "This sum is total opportunities" points to rows 1-13.
 - "This sum is total lost opportunities" points to rows 10-12.
 - "Averages of these are average throws %" points to rows 14-16.

Fig. 4: Player's performance and lost opportunities table

Turul	1,2	3,4	Final
DPMO	151133.5	120786.5	106339.5

Defects per million opportunities	Z
308537	2σ
66807	3σ
6210	4σ
233	5σ
3.4	6σ

Annotations in the diagram:
 - "We are here" points to the 3.4 value in the Z-table.
 - "We want to be here:" points to the 3.4 value in the Z-table.

Fig. 5: Performance level in sigma language

Example 2: The efficiency and behavior of a basketball player. Let us consider the case of a basketball player that performs during the season (4 stages and a playoff) as following (main performance parameters/total opportunities or games) (Fig. 4). By

opportunities we mean the particular situations during the game during which the athlete will exceed or fail (pass, throws, violations, fouls). The recorded data is from a basketball team performance measurements going back to 1990s.

What is the statistical meaning of these numbers?

Efficiency evaluation (σ) using the Efficiency Index (C.I.): C.I. (σ) = (Total number of opportunities-Lost Opportunities)/ (Total number of opportunities). In this case: C.I. (σ) = (39.7+35.6+48.9)-(6+4.3+5.2) / 41.4 = [(124.2/3)-(15.5/3)] / 41.4 = (41.4-5.17) / 41.4 = 0.88.

Total opportunities are the sum 1 through 13-14:

- Number of lost opportunities-sum of 8 though 12
- Efficiency for the phases 1-2 during regular season is: C.I. = (39.7-6) / 39.7 = 33.7 / 39.7 = 0.84

If the throws percentage is included, the index can be changed accordingly: C.I. = (σ + Free Throws Percentage + Field Goals Percentage) / 3 C.I. = 0.88 + (78+84+88)/3 + (48+55+53)/3 = 73.58.

The player's performance measured using six sigma (Fig. 5):

Table 1: Risk factors for a player's performance

	Probability	Impact	Total	Corrective actions
Wrong Pass	5	5	25	-develop strength and speed; -exercise passing in special situations-crowded, fast developing, faking, contro.
Missed in the paint	3	5	15	-field throws with contact and against aggressive defense
Missed free throws	3	5	15	-train and develop precision in fatigue and stress conditions
Missed fast break	3	5	15	-transition training
Fouls committed	3	3	9	-exercise aggressive defense and pressing man to man and zone
Missed 2 point field goals	3	3	9	-train throws in aggressive defense situations
Offensive fouls	1	5	5	- model game situations, train fast breaks
Technical	1	5	5	- same
Lost possession	1	3	3	-exercises for development of handling, technical dribble and possession
Travel	1	3	3	-synchronicity training
3 sec violation	1	1	1	-spatial and temporal awareness development
Missed 3 point field goal	1	1	1	-practice extensively 3 point field throws

Table 2: The convolution of the errors and their effects color coded by the gravity of the combination in basketball (5 is red = serious error, 3 is yellow = moderate error, 1 is green = minimal error)

	Offensive rebound	Defensive rebound	Intercept	Steal	Won 1x1
Wrong pass	5	5	3	3	3
Miss in the paint	5	3	3	3	1
Miss semi-distance	3	1	3	1	1
Miss field goal distance	3	1	1	1	1
Miss fast break	5	3	3	3	3
Miss free throw	5	3	3	3	3
Lost ball	5	3	3	1	1
3 sec violation	3	3	3	1	1
Offensive foul	5	5	3	3	3
Traveling	5	5	5	5	5
Foul	3	3	3	3	3
Technical foul	5	5	5	5	5
Flagrant Foul	5	3	5	5	3
Sentoffs	5	5	5	5	5

$DPMO = 1000000 \times (\text{lost opportunities}) / (\text{Total opportunities})$
 $DPMO (1,2) = 1000000 \times 6 / 39.7 = 151133.5$ (phases 1,2 regular season)
 $DPMO (3,4) = 1000000 \times 4.3 / 35.6 = 120786.5$ (phases 3,4 regular season)
 $DPMO (\text{playoff}) = 1000000 \times 5.2 / 48.9 = 106339.5$ (playoff)

Results show that the athlete performs at a level between 2σ - 3σ so we reduced the data to a statistical qualitative evaluation of the player in terms of DPMO (151133.5, 120786.5 and 106339.5, respectively) and we are targeting a 4 or better σ if possible (eventually until we reach 3.4 DPMO).

A risk analysis is now required (Table 1) and interpret the effect on each of the team's player as well as the interaction between the players. The coach is always concerned about the risks involved when making the team or changing a player during the game. A quality scorecard for each player could then become an important document describing not only the performance and the qualities or imperfections of the

player but also his interaction with the team and the other players. Only defects (negative effects) are emphasized in assessing the risks associated because efficiencies correlate with the total number of lost opportunities.

Making the latter vanish means achieving maximum efficiency. Statistically speaking, it doesn't matter whether the player scored 40 points per game alone, especially if his efficiency has been low, i.e. 20% vs. 50% (see efficiency index). A legend of the data listed in Table 2 is given below.

Legend:

- Probability 1, 3, 5 means low, medium and high risk levels. A total of 15-25 means a higher risk, 3-9 means medium risks and 3 means minimal risk
- Corrective actions column is perhaps the most important of all because it establishes the roadmap and strategies of correcting and reduction of the defects contained in the other columns
- The Fishbone diagrams (Fig. 6) show the risk factors need to be closely examined. Other elements can also be added to this diagram, such as team players, team functionality, the specific fitness and perhaps even the opponent team

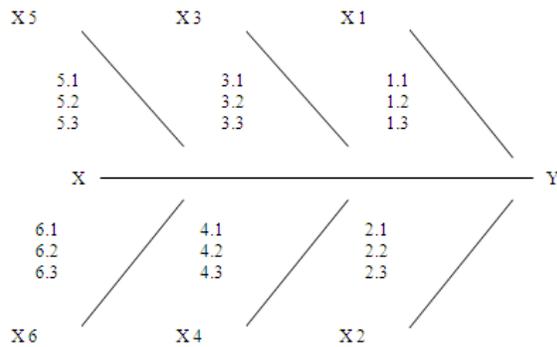


Fig. 6: Cause and effect diagram (Fishbone diagram)

The impact of defects or errors is also extremely important, since it marks the connected nature of chain errors that may lead to lowering performance, for instance a wrong pass after a defensive rebound, followed by an intercept by the opponent, (Table 2)

Note: It is essential to distinct between the level of error and the importance of the successes. An interception at the end of a game followed by a winner goal is more important than the same being achieved during the game. Conversely, a lost ball is much more decisive at the end of a close game than at the beginning of the game. Therefore the coach must have a clear reward evaluation system: (a) importance of the errors and (b) reward of successes.

CONCLUSION

- Six σ is the quality improvement method that can be successfully applied to coaching through objective assessment of data and efficient optimization of the training process. The success of the method depends heavily on the ability of the coach and trainer to record data, measure, interpret, analyze and control the data to correctly down select the right coaching method and determine the vital few elements that need special attention
- The examples in this study clearly show the feasibility of the method; they also demonstrate that a new methodology can be applied in a novel manner to positively influence the coaching system and introduce a new rationale in the process of coaching. It is worth mentioning that tools like the Fishbone diagram (failure analysis), efficiency index based on lost per total opportunities, DPMO, risk analysis and abatement, athlete's or team's score card and inter-relation between risk elements can and will improve dramatically the efficiency of the coaching process

- The method can be equally applied to individual and team sports. The athlete's performance can be estimated and improved upon at the individual level as well as team level
- A significant difference exists between the current, traditional approach of coaching and training and the statistically driven six sigma method. While traditionally coaching is based on the experience of the athlete and coach and historical events, six sigma determine course of action based on mathematically proven, statistically demonstrated and down selected reasons for poor performance based on data. Six sigma decisions are based on real data and systematically applied exactly to reduce and eliminate the root cause of large variance and poor performance. Six sigma methodology, when combined with coaching experience, allows a significant, data-driven improvement of the performance of an athlete or a team. The authors recommend the introduction of this validated method in coaching as a reliable, predictable tool of improving performance.

REFERENCES

- Ryan, B.F. and B.L. Joiner and J.D. Cryer, 2005. Minitab Handbook. 5th Edn., Cengage Learning, ISBN: 10: 0534496008, pp: 505.
- Breyfogle, W.F., 2003. Implementing Six Sigma-Smarter Solutions Using Statistical Methods. 2nd Edn., John Wiley and Sons, ISBN: 10: 0471265721, pp: 1229.
- Levine, D.M., P.P. Ramsey and K.R. Smidt, 2000. Applied Statistics For Engineers and Scientists. 1st Edn., Prentice Hall, ISBN: 10: 0134888014,
- Olugu, E.U and K.Y Wong, 2009. Supply Chain performance evaluation: Trends and challenges. Am. J. Eng. Applied Sci., 2:202-211. DOI: 10.3844/ajeassp.2009.202.211 URL: <http://www.thescipub.com/abstract/10.3844/ajeassp.2009.202.211>
- Pande, S. P., P.R. Neuman and R.R. Cavanaugh, 2000. The Six Sigma Way: How GE, Motorola, and Other Top Companies are Honing Their Performance. 1st Edn., McGraw-Hill, ISBN: 0071358064 (0-071-35806-4), pp: 448.
- StatSoft, Inc., 2000. Electronic Statistics Textbook. StatSoft, Tulsa, OK, USA. <http://sunsite.univie.ac.at/textbooks/statistics/stathome.html>