Teaching Statistical Thinking and Exploratory Data Analysis by Using the Data from the Red Bead Game

***Y. Adler, E. Hunuzidi, and V. Shper.* – MIS&A**
O. Maksimova - MPEI

"When planning for a year, plant corn.
When planning for a decade, plant trees.
When planning for life, train and educate people."
– Chinese proverb
"I hear and I forget, I see and I remember, I do and I understand"
Confucius 450bc

Introduction
There is no doubt that the most effective way to educate people is playing with them. There are many different games that were coined for teaching various strata of people to numerous number of statistical methods and procedures, see for example [1-3]. Among all of them the Red Bead Game (RBG) or the Red Bead Experiment (RBE) [4-7] and its derivatives [8-12] are probably most famous and widespread. In fact Dr. Deming used this game during his four-day seminars [4-6] for teaching the participants the main elements of his System of Profound Knowledge (SoPK) which consists of four components [5]:
Appreciation for a system
Knowledge about variation
Theory of knowledge
Psychology.
"The various segments of the system of profound knowledge can not be separated. They interact with each other" [5] and the knowledge of any one component is incomplete without others. That is why the game for teaching the SoPK should include all different facets of this notion. And the RBG possesses this ability. Speaking more simply Dr. Deming taught the listeners not the tools but an attitude and not crunching numbers but thinking before any calculations. This goes in line with the ideas of another statistical giant - John Tukey – who explained his notion of Exploratory Data Analysis (EDA) as follows (http://en.wikipedia.org/wiki/John_Tukey):

"If we need a short suggestion of what exploratory data analysis is, I would suggest that

1. It is an attitude AND
2. A flexibility AND
3. Some graph paper (or transparencies, or both).

No catalogue of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. The graph paper - and transparencies - are there, not as a technique, but rather as recognition that the picture-examining eye is the best finder we have of the wholly unanticipated."

And we think that the RBG let us teach the students both elements of the SoPK in a Deming's sense as well as elements of EDA in a Tukey's sense. That's why we consider that using the RBG for teaching students is so important.

1. The traditional RBG.

The traditional RBG is widely known as the experiment in which a "corporation" is formed from "willing workers", quality control personnel, a data recorder, and a foreman. The corporation's product are white beads, which are produced by dipping a paddle into a box of beads. Each paddle forms a batch from the production line. The paddle has 50 holes in it, and each hole will hold only one bead. Unfortunately, there are not only white beads in the box, but some amount of red beads – “defects”…The "willing workers" extract batches sequentially, inspector counts the number of defects, and a foreman demonstrates *all flaws of the current style of traditional management by objectives*: he praises for batches containing less "defects" than a pre-established number of red beads and blames and shames someone for exceeding that number and so on. The meaninglessness of such style of management is pretty clear due to simplicity of the RBG and from the evident fact that process' outcomes depend only on the system (the box's contents) do not depend on the efforts of "willing workers". And just this was the main goal of Dr. Deming. He was not focused on statistical tools themselves but on an understanding of Shewhart control chart as a tool for system analysis by using the basics of the theory of variability and psychology. As a result the Traditional RBG allows to demonstrate that

- *If the process is stable the system defines its outcomes and the process still has a lot of variations;*
- *Variations between batches and workers are caused by system and not by workers themselves;*
- *Anybody’s results today are useless for prediction tomorrow’s results;*
- *Ranking, pleading with workers, making performance appraisals, setting goals, etc. has no sense and only demoralizes “willing workers”;*
- *Mechanical sampling is not the same as random sampling.*

Besides The Traditional RBG let the participants to become acquainted with Shewhart Control Chart of *pn*-type and even to learn how the chart limits may be calculated from their own data. Fig.1 shows a typical example of such a chart.

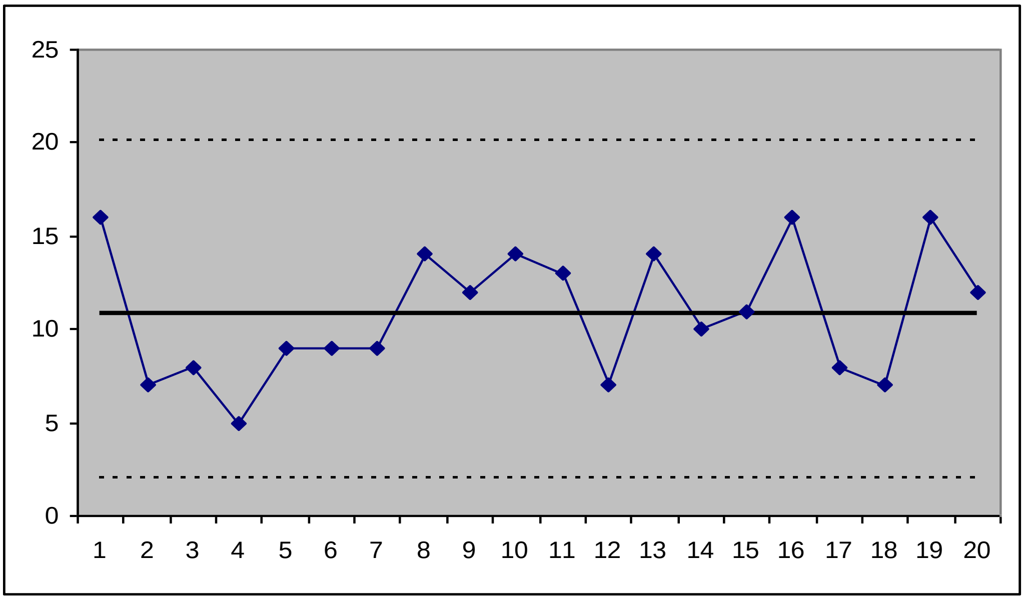


Fig.1

2. What changes have we made in our version of the RBG?
After many years of playing the RBG we came to conclusion that *this is an irreplaceable tool for teaching students both statistical thinking and EDA.* But to this end we have introduced some changes and additions into the game. Here they are.
1. After playing “first week” traditionally with four "willing workers" we stop returning red beads into the box and put them into a different box (i.e. we refuse from sampling with replacement and continue playing without replacement);
2. We introduce one more inspector named time-keeper who records the times of each batch preparation (= paddle extraction) for all workers (i.e. we get from our game not only data for counts but variable data as well).

Typical results for a number of red beads are shown in fig.2:

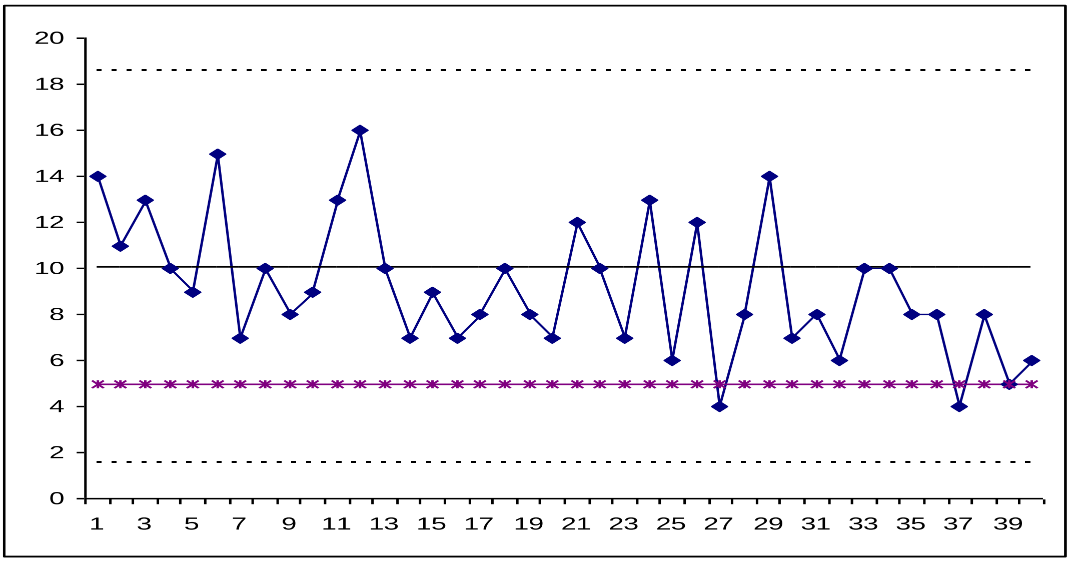


Fig.2

Typical results for sample extraction data (in seconds) are shown in fig.3:

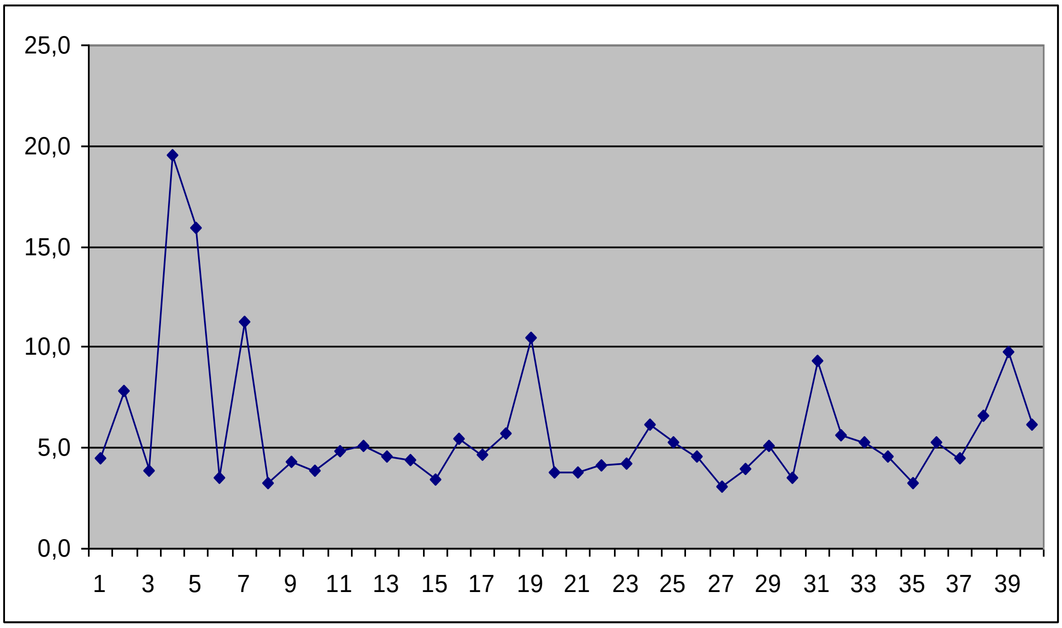


Fig.3

The marked line in fig.2 shows the goal: the number of red beads within each batch should not be more than five (this is our choice for the goal which let us usually have 2-4 lucky students who get their reward). We also use monetary and sweet rewards for increasing the fun while playing.

Thus we have two sets of data:

• 40 points of qualitative data (the number of red beads for each sample of 50 beads in a paddle);

• 40 points of quantitative data (the seconds for extracting each sample from a box).

What can we extract from these data in addition to all mentioned in ***1.***?

3. Additional information that follows from our version of the RBG?

First of all it is worth noting that we use all points listed in ***1.*** to explain students the basics of the SoPK: system approach to the variability of any process with the simultaneous attention to human psychology. Then we start teaching how to create and analyze the *pn*-chart. Here due to peculiarities of our data we can explain students the importance of reasonable choice of data used for control chart limits calculations. In fact all second week we permanently intervene into our system because the percentage of red beads inside the box is changing after each paddle drawing. So it is unreasonable to use the data from the second week for limit calculations: we need control limits in order to find out the limits of variation for a stable process when nothing intervenes into the process and we know that we do intervene… Besides we explain that reasonable use of the power of Shewhart Control Charts is based on the two phase procedure:

- in phase one – retrospective – we are trying to understand the process and assess its stability; and

- in phase two – prospective or monitoring – we are interested to predict future state of the process or monitor its behavior.

Different behavior of our system during the first and second weeks let students to grasp the importance of phase one and two for estimating the control limits.

Having taught students to calculate the limits for *pn*-chart we transfer to our time data which we use for constructing *x-mR* chart. Here we first of all return to the problem of the phase I and discuss whether we can use all data for estimating the limits or we should take only the first week data again. Usually most of the students quickly understand that this time the situation is quite different because from the viewpoint of operation procedure of time measuring both weeks are totally identical[1]. Then we explain how to create a *x-mR* chart and ask our students to look at our data more attentively. One can see obvious abnormalities in fig.3 – at some points the durability of operation was much more than at other points shown.

Why is this?

As a rule many students quickly remember that these were the points when our "willing worker" could not draw the wholly filled paddle and several holes remained empty. In this case "the willing worker" should (by the rules of the game) turn the paddle over and make the second, and sometimes the third attempt. But this means that the worker had made not one batch of the product but two or even three. And everybody knows that production of two batches will require twice as much time as for one batch, production of three batches – thrice as much time and so on. In other words these data are not homogenous and this is very important element of Shewhart control chart construction: analysis for homogeneity should precede any

[1] We'd like to note in passing that such a discussion of the same question but with different data sets is a tacit embodiment of the cycle of knowledge

calculations [13]. This let us explain students the necessity for understanding data before any crunching the numbers by formulas – it is useless to estimate the control limits by using all data as they describe different processes (drawing one or two or three … batches out of the process).
Thus by this moment we taught the students how to construct charts for count data (*pn*-chart and *p*-chart) and charts for measurements (*x*-*mR* chart and $\overline{X} - R$ charts).
What else can we learn from our data?
4. Teaching the students the basics of Exploratory Data Analysis (EDA)
First we ask our students to plot a histogram for the data of red beads.
Usually students have problems with discrete data even if they are able to plot a histogram for continuous data. After this we ask them to plot a binomial then Poisson and then Gauss distribution on their histogram. Here they must be able to find out the parameters of bin- and poi- models from their RBG data and to transfer them into parameters of normal distribution (by two ways). This is a very important moment because before mechanical substitution into formulas a bit of thinking is required: what are the *x* and *p* and *n* in this or that formula and where can I take these values from?
An example of picture we want to see at the screens of our students computers is shown in fig.4.

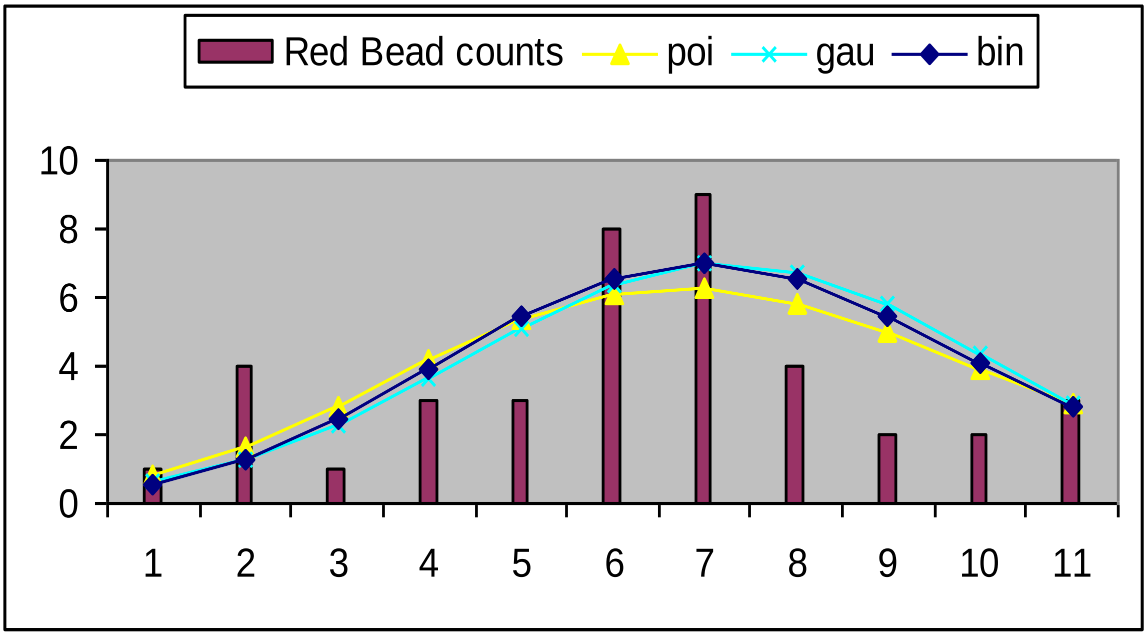
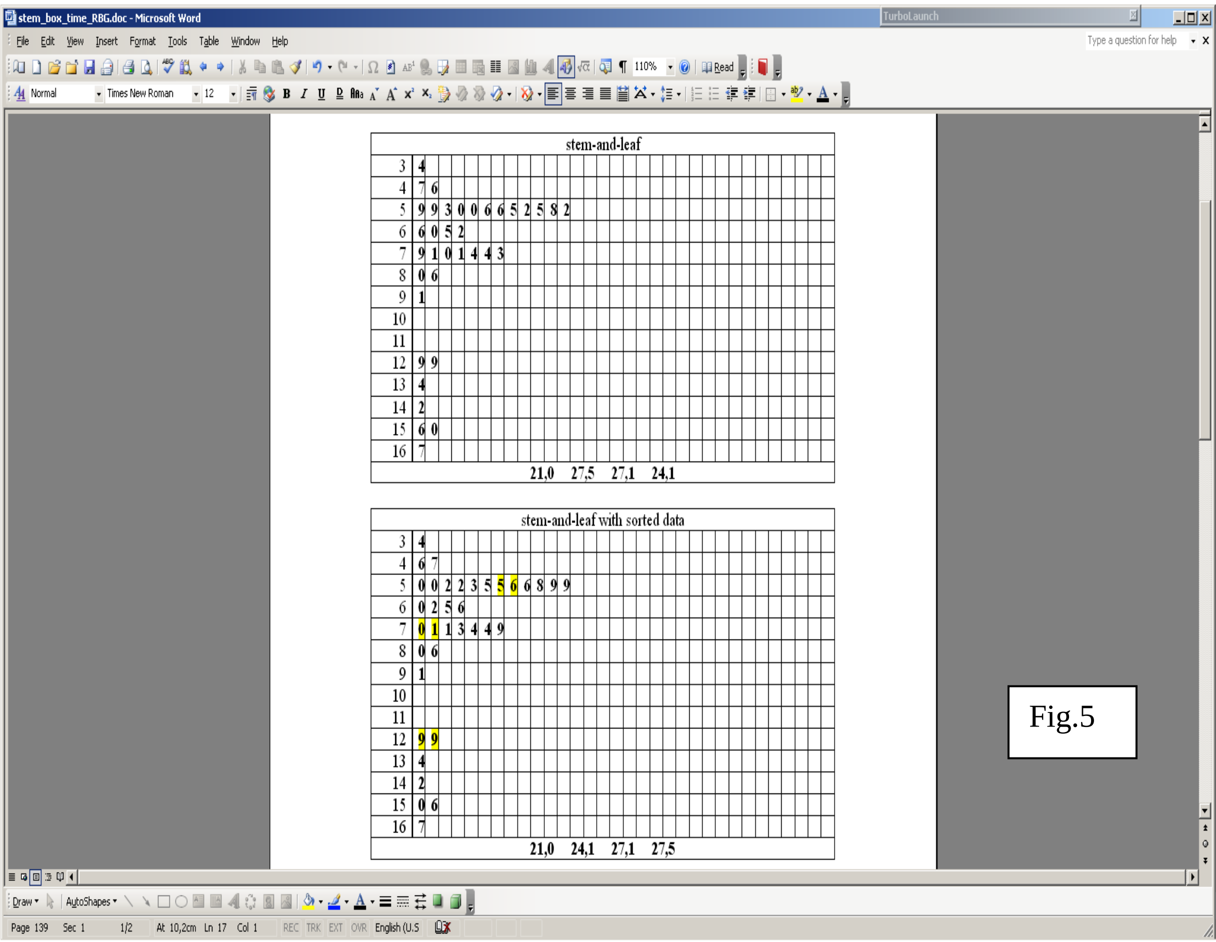


Fig.4

Then we take time data and teach the students to construct stem-and-leaf plot: first unordered and then arranged in ascending (e.g.) order. Simultaneously they are being acquainted with median and quartiles which are necessary for box-and-whisker plot and are learning to find out these values by hand and by using "quartile" function in Excel. Students make their handy calculations on a special paper form we prepare for them beforehand (see fig.5).
Then quite naturally we teach students how to construct a box plot. To this end we divide class into four groups (in accordance with our number of "willing workers") and they use data for each “willing worker” (ten points for each operator) for constructing the box plot and comparing the worker’s work. Here we explain the main advantages of box-and-whisker plot for comparing different samples, for revealing outliers and for

working with limited number of data. And again we ask students to plot the boxes by hand and by using Excel (see fig.6).



stem-and-leaf

3	4											
4	7	6										
5	9	9	3	0	0	6	6	5	2	5	8	2
6	6	0	5	2								
7	9	1	0	1	4	4	3					
8	0	6										
9	1											
10												
11												
12	9	9										
13	4											
14	2											
15	6	0										
16	7											

21,0 27,5 27,1 24,1

stem-and-leaf with sorted data

3	4											
4	6	7										
5	0	0	2	2	3	5	5	6	6	8	9	9
6	0	2	5	6								
7	0	1	1	3	4	4	9					
8	0	6										
9	1											
10												
11												
12	9	9										
13	4											
14	2											
15	0	6										
16	7											

21,0 24,1 27,1 27,5

Fig.5

(The horizontal axis in fig.6 presents student's names written in Russian. Two names for one box mean that one "willing worker" was fired for "bad performance" and a new worker continued playing).

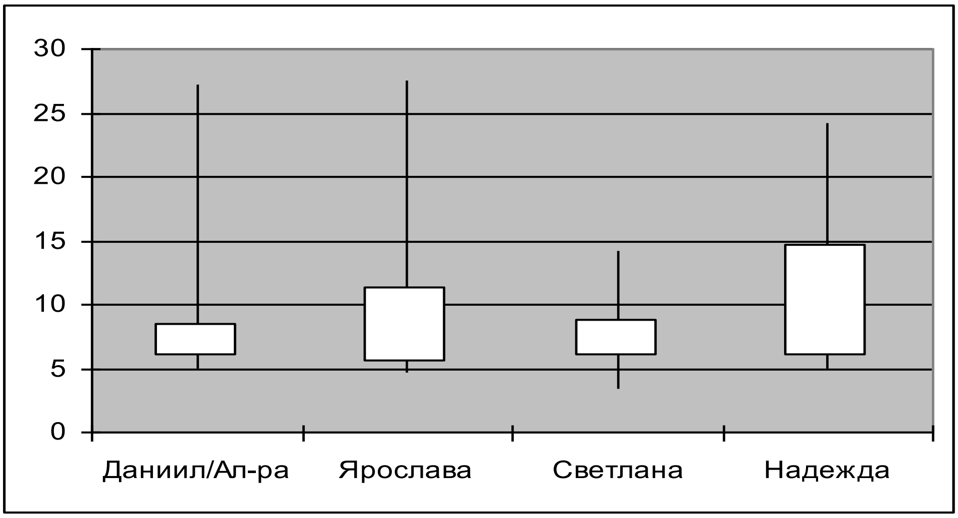


Fig.6

What else can we learn from our data?

Further we use the time data for teaching students how to construct the probability plot for different types of distribution (gau, lgau and wei) by using probability paper and how to interpret the graph obtained. A typical graph of one of our recent games is shown in fig.7. Again they must be able to do this by hand and by using Excel[2].

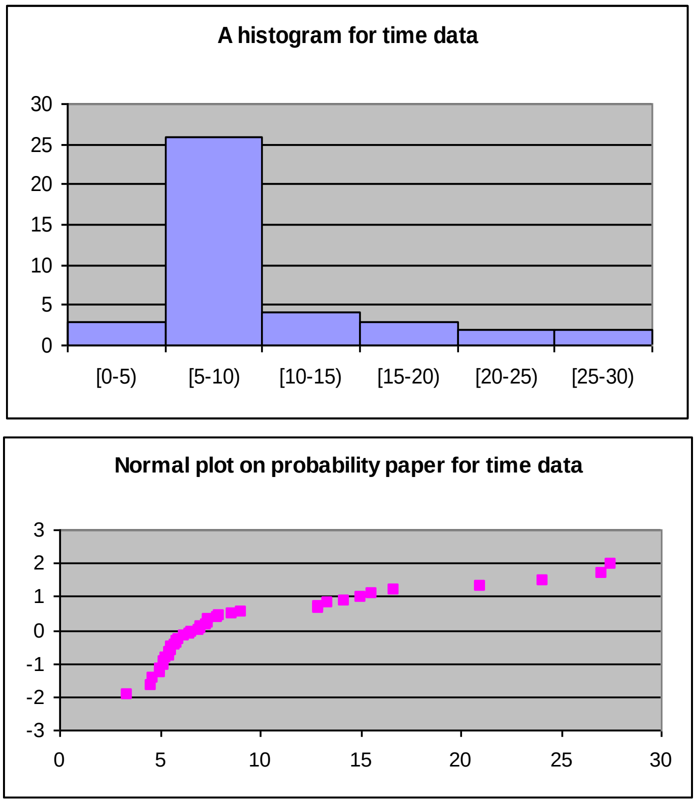


Fig.7

Some of the following items are optional – it depends on the level of our students and time for SPC course (usually this is 17 two-hours lectures and 34 two-hours practices during one term).

So when it is possible we explain our students the difference between different ways to estimate the standard deviation. For example for the data shown above one can obtain:

Sigma $_{(n-1)}$=6,24 sec

Sigma $_{(R/d2)}$=5,92 sec

This 5%-difference may turn out to be important for the values of control limits…i.e. for the definition of whether the process is stable or not.

Sometimes we discuss the capability indices and how to use them rightly. To do this we take our time data and establish some value of time for operation of drawing the beads (ten seconds for example).

By taking each day data as a rational subgroup it is possible to plot an average and range charts and discuss the problems of this chart plotting and interpretation.

At last we almost always teach our students to construct Cusum chart and discuss its interpretation. An example of such a chart is shown in fig.8.

[2] In fact students of our institute are being taught to work with StatGraphics package but we intentionally avoid using specialized packages at a tyro stage. Sometimes we suggest advanced students to plot all above-mentioned in any statistical software.

v

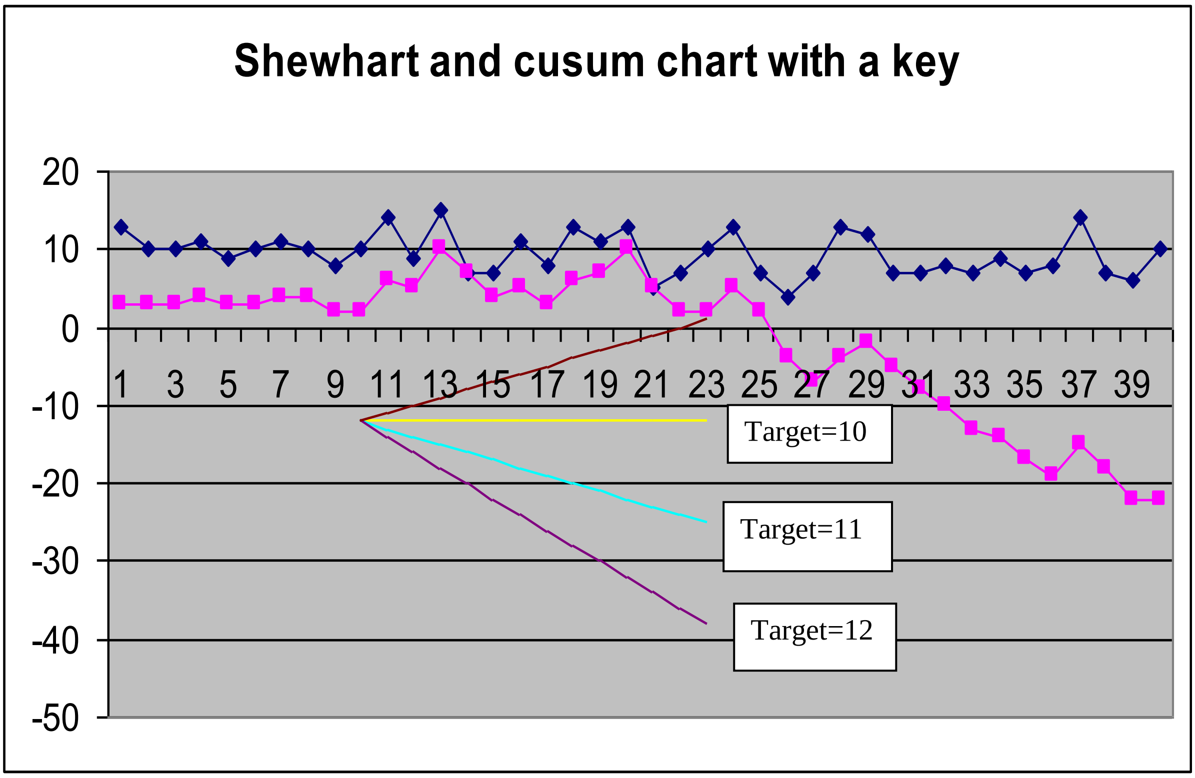


Fig.8

Besides we have two or three classes in one course (one term) – so it is possible to make comparisons between different batches

5. What are the main advantages of this approach ?

1. The students themselves generate the data which they completely understand, so that they can explain different nuances of these data and how these nuances can influence the results of data treatment.
2. The students study the basics of the SoPK and try to apply the elements of this system to the data they generated.
3. The students learn to work with one set of data by using many different tools and methods – this is very important for drawing useful information out of data…and for growth of their knowledge.
4. The RBG (RBE) is a very simple but interactive game – this creates quite positive atmosphere at the course from its very beginning…Many our students who have graduated from the institute several years ago remember the RBG very well.

Conclusion

In Wikipedia the article about EDA starts as follows: "In statistics, **exploratory data analysis (EDA)** is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. A statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modeling or hypothesis testing task" - http://en.wikipedia.org/wiki/Exploratory_Data_Analysis

We think that our version of teaching students by using the RBG deserves being widely spread as a way for teaching "… seeing what data can tell us" and simultaneously a way to spread the basics of statistical thinking among the students of practically all specialties.

References

1. **Morrow C., McNeese L.O**. A Simple Way to Digest SPC. – ASQC 56th Annual Quality Congress Proc., 2002, p.257.
2. http://www.onlinemathlearning.com/statistics-games.html
3. http://www.stats.govt.nz/tools_and_services/schools_corner/activities/interactive-games.aspx
4. **W.E. Deming.** Out of Crisis. – The MIT Press, 2000.
5. **W.E. Deming**. The New Economics for Industry, Government, Education. – The MIT Press, 1993.
6. **H. Neave**. The Deming Dimension. – SPC Press, 1990.
7. **S. Prevett**. Video on RBG - *www.youtube.com/view_play_list?p=8E522DD542C4CA69*
8. **Umble E.J., Umble M.M**. Focus On The Classroom. Developing Control Charts and Illustrating Type I and Type II Errors. – Quality Management Journal, 2000, v.7, no.4.
9. **Balestracci D.** An Alternative to the Red Bead Experiment. - http://www.qualitydigest.com/aug06/departments/spc_guide.shtml
10. **Olsen, T.** Deming's Quality Experiments Revisited. - 2007, *INFORMS Trans. Ed.* **8**(1) 37-40. Available online at http://www.informs.org/site/ITE/article.php?id=114
11. **Turner R.** The Red Bead Experiment for Educators. – QP, June 1998, p.69-74.
12. **Moore St.** Deming's Famous Red Bead Experiment With A Twist. - http://www.qualitydigest.com/inside/twitter-ed/deming-s-famous-red-bead-experiment-twist.html#
13. **Wheeler D.** The Four Questions of Data Analysis. - http://www.qualitydigest.com/inside/twitter-ed/four-questions-data-analysis.html