

p Charts on Excel

p Charts can be used when the subgroup size remains constant or when the subgroup size is varying. An example of each will be given.

Subgroup Size Remaining Constant

For setting up the np chart we used the following situation:

Suppose that a course has 50 students, and roll is taken at each class meeting. The number of students absent is recorded, and the following data are obtained for the first 15 class meetings: 4, 1, 3, 1, 2, 6, 3, 2, 3, 0, 12, 8, 7, 6, 6.

We could have set up a p chart instead of an np chart to analyze this process. To do this, we would convert each of the observations from a count of the number absent to a proportion absent. This is accomplished by dividing each of the counts above by 50 ($=\text{cellreference}/50$). We can use a formula to create an additional column in Excel with these values. The new spreadsheet will look like the following:

	A	B	C	D
1	Class	Absent	proportion	
2	1	4	0.08	
3	2	1	0.02	
4	3	3	0.06	
5	4	1	0.02	
6	5	2	0.04	
7	6	6	0.12	
8	7	3	0.06	
9	8	2	0.04	
10	9	3	0.06	
11	10	0	0	
12	11	12	0.24	
13	12	8	0.16	
14	13	7	0.14	
15	14	6	0.12	
16	15	6	0.12	

The values in Column C provide the data for plotting the run chart that will become the control chart. To calculate the control limits we will need to find the average proportion absent. This is done by taking a ratio of the total number absent to the total number possible (in this case 50 per day for 15 days or 750). To find the total number absent, we use the formula =sum(B2:B16). Therefore, the average proportion absent is $64/750 = .085333$ (found using =B18/B19). To calculate limits we use functions and the formulas for the [p chart](#) with constant subgroup size. The relevant part of the spreadsheet is shown below:

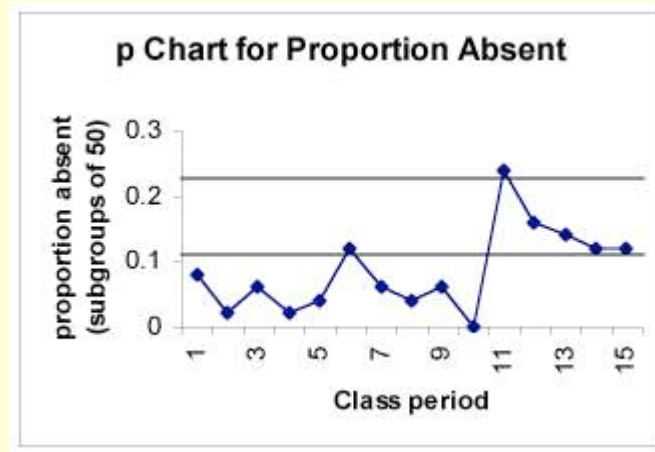
	A	B	C
9	8	2	0.04
10	9	3	0.06
11	10	0	0
12	11	12	0.24
13	12	8	0.16
14	13	7	0.14
15	14	6	0.12
16	15	6	0.12
17			
18	total absent	64	
19	total possible	750	
20	average proportion		0.085333
21	LCL		-0.0332
22	UCL		0.203863
23			
24			

The formulas entered are listed below:

Cell Reference	Formula	Notes
B18	=SUM(B2:B16)	Finds the total number absent.
B19	=15*50	Finds the total number of absences that could have been recorded.
C20 (calculates the CL)	=B18/B19	Finds the average proportion absent. This figure is used any place p-bar is needed.

C21 (calculates the LCL)	=C20-3*SQRT(C20*(1-C20)/50)	Substitutes into the control chart formulas and recognizes that the subgroup size is 50.
C22 (calculates the UCL)	=C20+3*SQRT(C20*(1-C20)/50)	

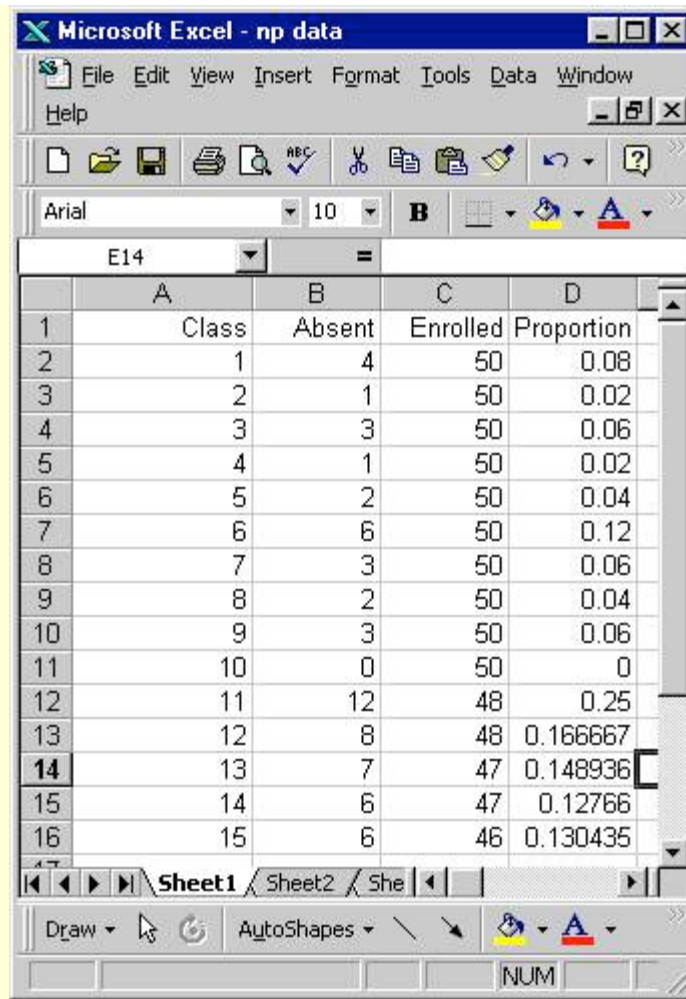
Plotting the run chart and adding the control limits provides the following chart:



We should not be surprised to find the same signal on this chart that we found on the np chart—after all we are asking the same question about the same process.

Subgroup Size Varying

Now, suppose that we learn that the number of students did not remain constant. In fact, some students dropped the course following the first test. To accommodate this we will change Column C to indicate the number of students registered for the course. Then we will calculate the proportion absent in Column D. Each entry in Column D will be obtained by dividing the entry in the corresponding row of Column B by the entry in Column C. The resulting spreadsheet is shown below:

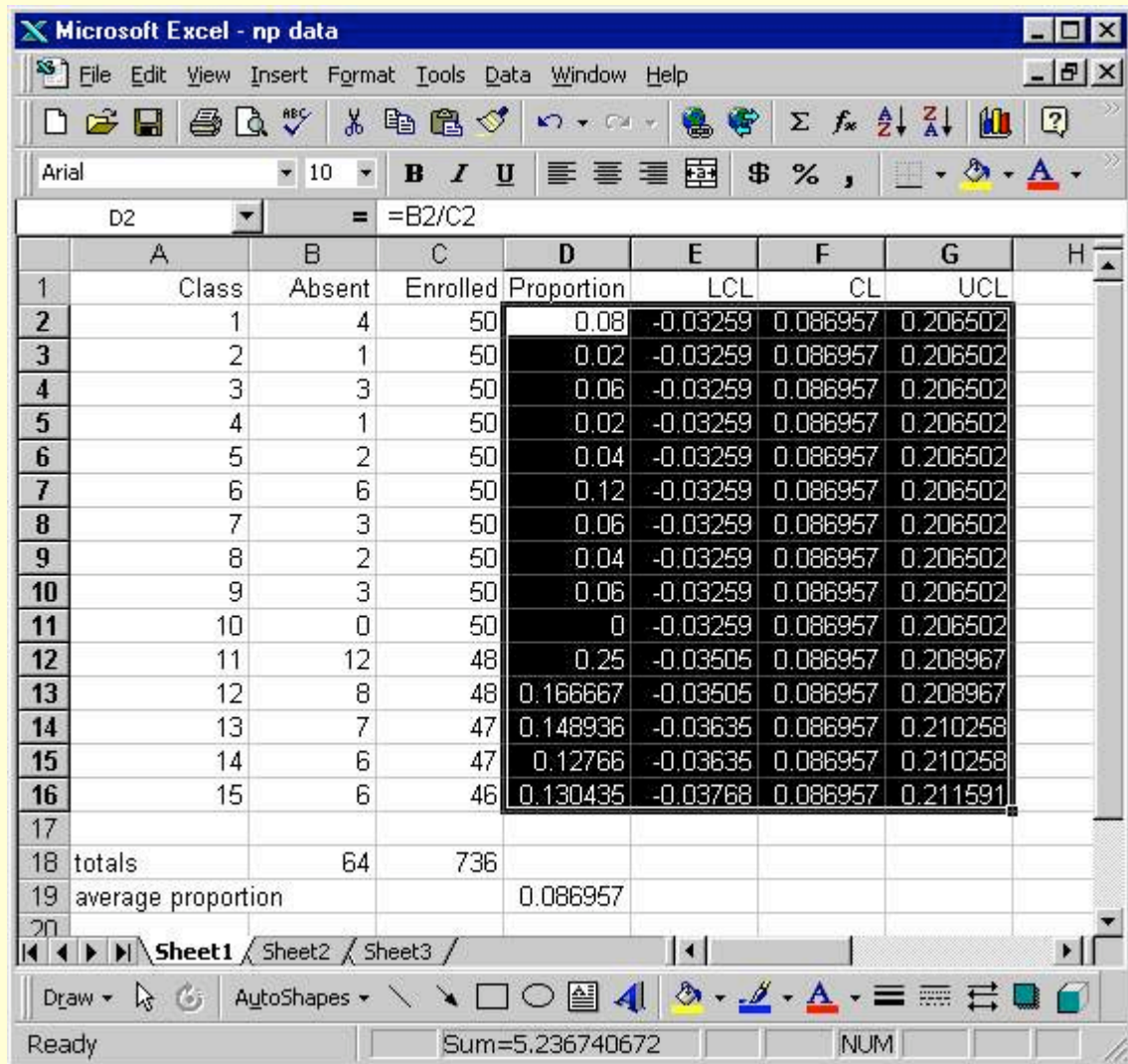


Column D now provides the data for plotting the dots on the run chart. Calculating limits for a [p chart](#) is slightly more complicated when the subgroup size does not remain constant. The formulas require that we adjust the limits to accommodate the varying sample size. Therefore, separate limits must be calculated and plotted for each point. To do this, we create three more columns in the spreadsheet—one for the LCL, one for the CL, and one for the UCL. The center line will remain constant at a value of \bar{p} , so we will calculate that quantity first. To do this, we find the total number of absences [=sum(B2:B16)] and divide this by the total number of possible absences [=sum(C2:C16)]. For this example, the result is in D19. We will create an absolute reference to this cell for the entries in the column labeled CL [=D\$19]. To calculate control limits, we will use functions and formulas for the entries that correspond to the first day. Then we will copy these down the column. The formulas used are listed below:

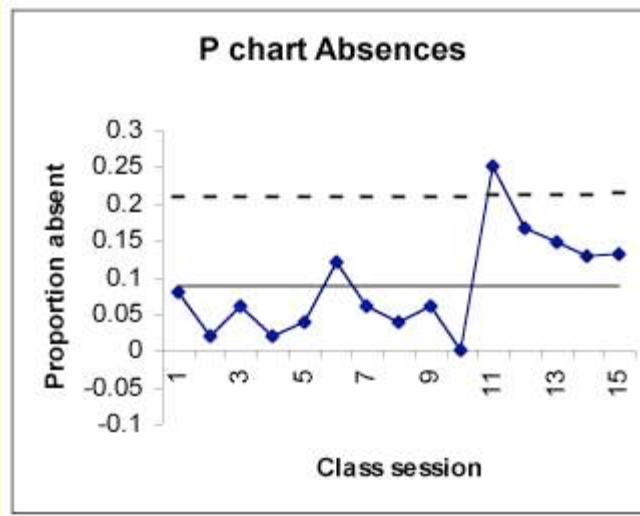
Cell Reference	Formula	Notes
F2 (calculates the CL)	\$D\$19	This creates a column with \bar{p} as the entry in each row.

E2 (calculates the LCL)	=F2-3*SQRT(F2*(1-F2)/C2)	This references the subgroup size (from Column C) in the formula.
G2 (calculates the UCL)	=F2+3*SQRT(F2*(1-F2)/C2)	

Once the entries in these three columns are complete for this row, the other rows in the columns can be completed by copying or filling down the column. The resulting spreadsheet is shown below:



Once this is done, the control chart can be plotted by selecting the data highlighted above. Once the line chart is created, some clean-up is still in order. The center line needs to be changed to a solid line, the control limits and center line all need to be shown in the same color (actually, the lower control limit does not need to be shown for this problem since the calculated values are negative and negative numbers do not have meaning for this situation), and the control limits need to be shown as markers above/below each point rather than as lines. If you look carefully at the plot (or at the data), you will see that the limits do change as the subgroup size changes. The final control chart is shown below:



From this plot we still see a signal on day 12. Further investigation revealed that this was the day following the test!

Back to [Calculating Control Limits with Excel](#)

Back to [Excel Hints Page](#)

Back to [BUSA 3110 Home Page](#)