The T Chart in Minitab Statistical Software

Background

The T chart is a control chart used to monitor the amount of time between adverse events, where time is measured on a continuous scale. The T chart is an extension of the G chart, which typically plots the number of days between events or the number of opportunities between events, where either value is measured on a discrete scale. Like the G chart, the T chart is used to detect changes in the rate at which the adverse event occurs.

When reading the T chart, keep in mind that points *above* the upper control limit indicate that the amount of time between the events has increased and thus the rate of the events has *decreased*. Points *below* the lower control limit indicate that the rate of adverse events has *increased*.

The Transformation Approach

The T chart is included in other software packages, all of which transform the data for time between events to make it more normally distributed. The transformed data are used to determine the control limits, which are then converted back to the original data scale and plotted with the original data.

The problem with this approach is that the tails of the transformed data do not fit a normal distribution very well. With the transformation approach, the probability of a point being outside the control limits is only 0.0007546. In contrast, with a standard control chart based on a normal distribution (such as an I chart or an Xbar chart), the probability of a point being outside the control limits is much higher, 0.00269. The transformation method for a T chart results in an unusually low probability of out-of-control points and thus an inflated Average Run Length (ARL).

Simulations (see Table 3 below) show that the false alarm rate increases exponentially for extremely skewed data and decreases to almost 0 for data that are less skewed. In general, a T chart implemented with the transformation approach has very low detection capability, especially at the lower control limit. The low power at the lower control limit means that the chart has virtually no ability to detect increases in the adverse event rate.

The Exponential Distribution Approach

Another approach to the T chart is to model the time between events using an exponential distribution. The basis for this model is that, if adverse events occur according to a Poisson model, then the time between events should follow an exponential distribution. This approach uses percentiles of the exponential distribution corresponding to the ± 1, 2, and 3 sigma zones in a standard chart based on the normal distribution. These percentiles are sometimes called “probability limits”. The use of probability limits means two things:

- The ARL and false alarm rate for an in-control process are the same as one would expect in an Xbar chart with normally distributed data.
- The expected ARL and false alarm rates apply only if the data are from an exponential distribution.

The issue with the exponential distribution is that, although it is the theoretically correct distribution for time between Poisson events, the data in practice often follow a slightly different model. The data may appear to be exponentially distributed, but may actually deviate enough to seriously impact the ARL and false alarm rate. If the data come from a distribution that is more skewed than an exponential distribution, the false alarm rate can be extremely high at the lower limit, meaning that there would be a high incidence of falsely concluding that the adverse event rate had increased. On the other hand, if the data come from a distribution that is less skewed than an exponential distribution, the power to detect increases in the adverse event rate goes to 0.

The exponential distribution has a skewness value of 2 and a kurtosis value of 6. Simulations (see Tables 1 to 3 below) show that, as the skewness and kurtosis of the data increase from these values, the false alarm rate associated with the lower control limit increases exponentially. The false alarm rate associated with the upper control limit increases more slowly. As the skewness and kurtosis of the distribution decrease from the exponential values of 2 and 6, the false alarm rate associated with the upper control limit increases, while the false alarm rate associated with the lower control limit goes to 0.

Minitab's Approach—the Weibull Distribution

In order to increase the robustness of the chart, Minitab uses a Weibull distribution rather than an exponential distribution to model the time between events. The Weibull distribution has 2 parameters, shape and scale. If the shape parameter is equal to 1, the Weibull distribution is the same as an exponential distribution with the same scale parameter as the Weibull distribution.

Varying the shape parameter around 1 allows the Weibull distribution to take on many different shapes, from extremely peaked and extremely right skewed (for a shape parameter of less than 1), to symmetric (for a shape parameter of about 3), to left skewed (typically for shape parameter greater than 5). It is expected that the shape parameter will typically be between 0.5 and 2, because the distribution would then be close to the expected exponential distribution. Although using probability limits from a Weibull distribution still means that the expected ARL and false alarm rate would only apply if the data are in fact from a Weibull distribution, this broader family of distributions will increase the chances of obtaining a good fit.

Simulations

For the following tables, 100 random samples of 10,000 data points each were simulated from the specified distribution. The proportion of points outside the control limits is shown in the table. For a standard chart based on the normal distribution, such as an Xbar chart, the expected proportion of points outside the limits is 0.00269.

The simulations use the Weibull and chi-square distributions. A chi-square distribution with 2 degrees of freedom is the same as an exponential distribution with a mean of 2. Varying the degrees of freedom around 2 makes the chi-square more or less skewed than an exponential. See Figure 1. A Weibull distribution with a shape parameter of 1 is the same as an exponential distribution with a mean equal to the scale parameter from the Weibull distribution. Varying the shape parameter around 1 makes the Weibull more or less skewed than an exponential. See Figure 2.

Table 1a: Exponential-based T chart with chi-square data

Sampling from a Chi-Square Distribution						
Degrees of freedom	**Skewness**	**Kurtosis**	**Below LCL**	**Above UCL**	**Total Outside Limits**	**% of Expected Outside**
0.5	4.03	24.21	0.149413	0.027207	0.17662	6565.80%
0.75	3.34	17.31	0.065271	0.016747	0.08218	3055.02%
1.0	2.81	11.73	0.029265	0.010133	0.039398	1464.61%
1.25	2.55	9.9	0.013345	0.006102	0.019453	723.16%
1.5	2.3	7.83	0.006159	0.003682	0.009841	365.84%
1.75	2.12	6.63	0.002868	0.002223	0.005091	189.26%
2.0	1.96	5.85	0.001351	0.001339	0.00269	100.00%
2.5	1.8	4.9	0.000227	0.001839	0.002689	99.96%
3.0	1.64	4.03	0.000037	0.004179	0.004216	156.73%
3.5	1.52	3.39	0.000006	0.006696	0.006702	249.14%

Table 1b: Exponential-based T chart with Weibull data

Sampling from a Weibull Distribution						
Shape	Skewness	Kurtosis	Below LCL	Above UCL	Total Outside Limits	% of Expected Outside
2.0	0.63	0.25	0.000002	0	0.000002	0.07%
1.75	0.83	0.69	0.00001	0	0.00001	0.37%
1.5	1.07	1.4	0.000049	0	0.000049	1.82%
1.25	1.43	2.77	0.000258	0.000025	0.000283	10.52%
1.0	2.02	6.13	0.001335	0.001355	0.00269	100.00%
0.75	3.08	15.04	0.007002	0.016195	0.023197	862.34%
0.5	6.32	68.71	0.036022	0.076449	0.112461	4180.71%

Table 2a: Weibull-based T chart with chi-square data

Sampling from a Chi-Square Distribution						
Degrees of freedom	Skewness	Kurtosis	Below LCL	Above UCL	Total Outside Limits	% of Expected Outside
0.5	4.03	24.21	0.006500	0.000000	0.006288	233.75%
0.75	3.34	17.31	0.004558	0.000013	0.004571	169.93%
1.0	2.81	11.73	0.003635	0.000115	0.003750	139.41%
1.25	2.55	9.9	0.002788	0.000506	0.003294	122.45%
1.5	2.3	7.83	0.001945	0.000621	0.002566	95.39%
1.75	2.12	6.63	0.001656	0.000898	0.002554	94.94%
2.0	1.96	5.85	0.001596	0.001529	0.003125	116.17%
2.5	1.8	4.9	0.000877	0.001849	0.002726	101.34%
3.0	1.64	4.03	0.000585	0.002055	0.002640	98.14%
3.5	1.52	3.39	0.000453	0.003190	0.003643	135.43%

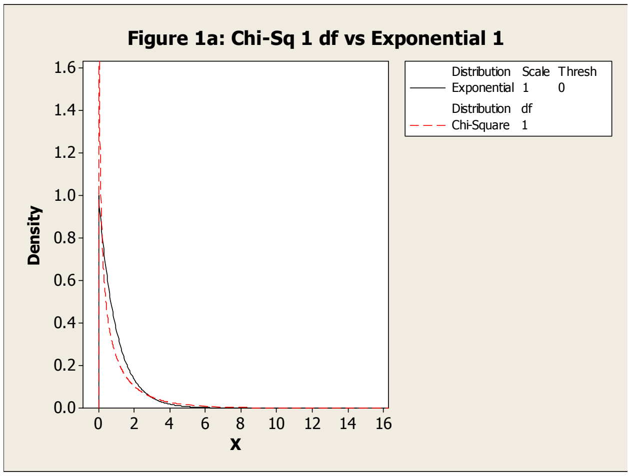
Table 2b: Weibull-based T chart with Weibull data

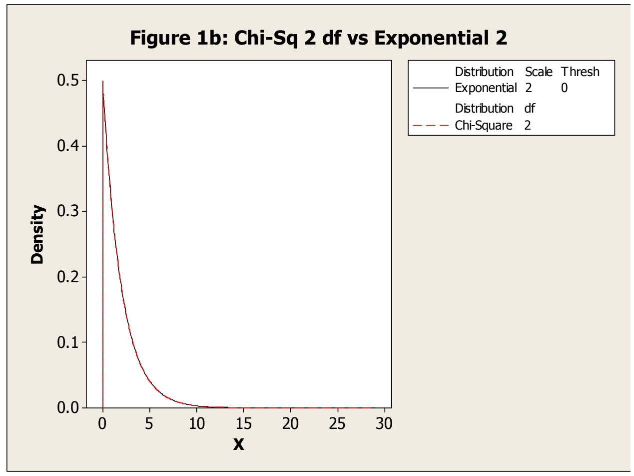
Sampling from a Weibull Distribution						
Shape	Skewness	Kurtosis	Below LCL	Above UCL	Total Outside Limits	% of Expected Outside
2.0	0.63	0.25	0.001386	0.001538	0.002924	108.70%
1.75	0.83	0.69	0.001365	0.001521	0.002886	107.29%
1.5	1.07	1.4	0.001335	0.001312	0.002647	98.40%
1.25	1.43	2.77	0.001299	0.001351	0.002650	98.51%
1.0	2.02	6.13	0.001475	0.001498	0.002973	110.52%
0.75	3.08	15.04	0.001368	0.001347	0.002715	100.93%
0.5	6.32	68.71	0.001220	0.001375	0.002595	96.47%

Table 3: Transformation-based T chart with chi-square data.

Sampling from a Chi-Square Distribution				
Degrees of freedom	Skewness	Kurtosis	Total Outside Limits	% of Expected Outside
0.5	4.03	24.21	0.14365	5340.15%
0.75	3.34	17.31	0.073075	2716.54%
1.0	2.81	11.73	0.0367875	1367.57%
1.25	2.55	9.9	0.017825	662.64%
1.5	2.3	7.83	0.0095125	353.62%
1.75	2.12	6.63	0.00523333	194.55%
2.0	1.96	5.85	0.00256	95.17%
2.5	1.8	4.9	0.00069333	25.77%
3.0	1.64	4.03	0.00034	12.64%
3.5	1.52	3.39	0.00017333	6.44%

Figure 1: Comparing chi-square and exponential distributions





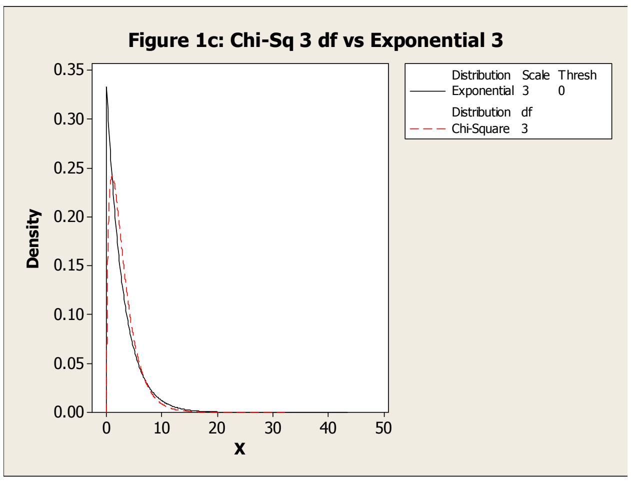
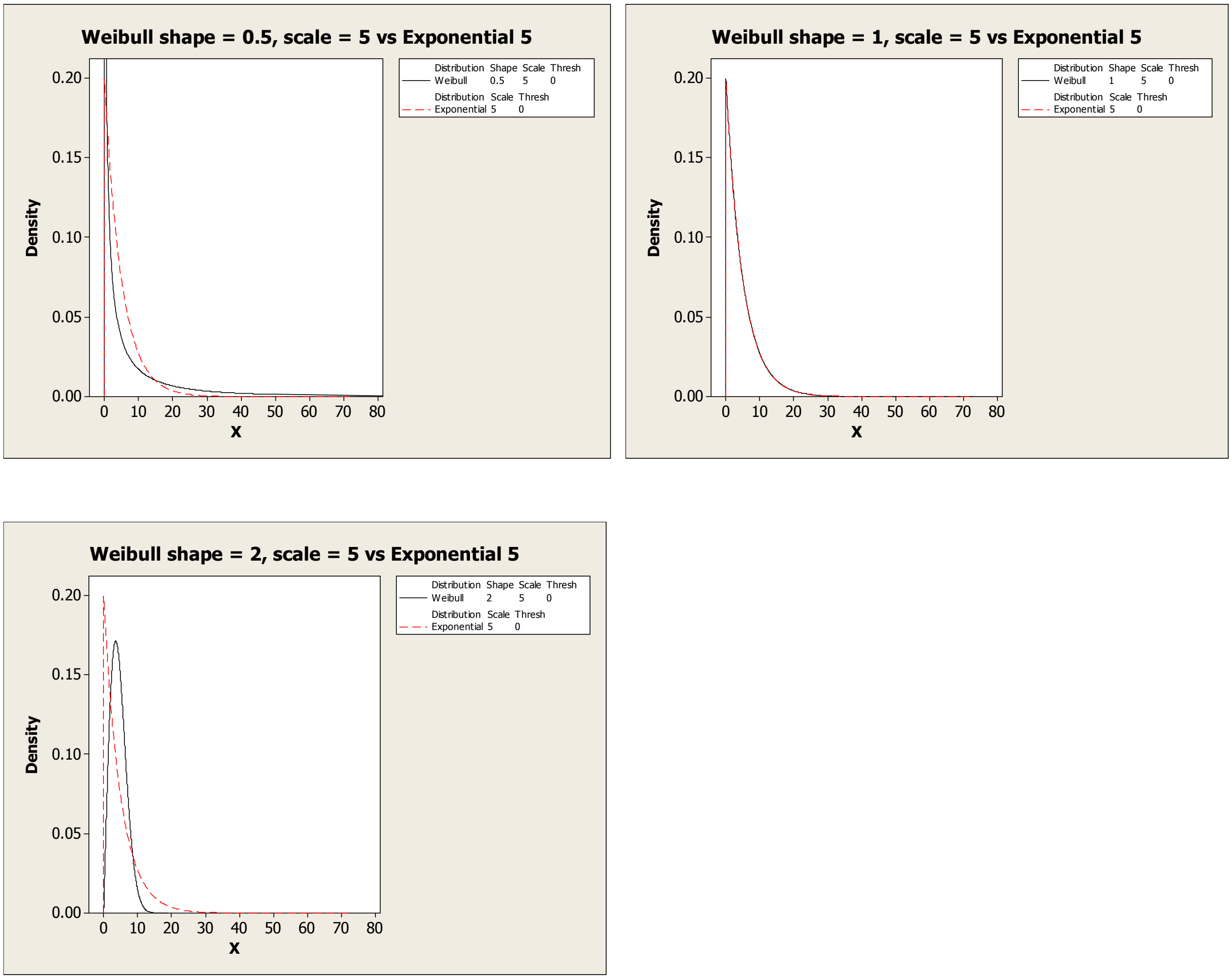


Figure 2: Comparing Weibull and exponential distributions



Use Cases for the T Chart

The difference between a G chart and a T chart is the scale used to measure distance between events. The G chart uses a discrete scale (counts of days between events or opportunities between events recorded as integers). The T chart uses a continuous scale (usually the dates *and* times that the events occurred). Most uses of the T chart discussed in research are about monitoring infection rates in healthcare settings. Other examples include monitoring medication errors, patient falls and slips, surgical complications, and other adverse events.

Note that it is not necessary to have both dates and times. In fact, it is expected that a prominent use case will be having date-only data. If the number of opportunities per day is not relatively constant, then a T chart may be a better choice than a G chart.

Properties of the T Chart

Like other control charts, the T chart has a center line and upper and lower control limits. There are also zones corresponding to the ± 1, 2, 3 sigma zones in an Xbar chart or an I chart. These zones are not displayed in the chart, but they are used in the tests for special causes. The control limits and zones are all based on percentiles of the Weibull distribution. They are not multiples of the standard deviation above and below the center line, as in other charts. As a result, the control limits and zones are not symmetric around the center line, except in the rare case where the Weibull distribution itself is symmetric.

The data that are plotted on the chart are the number of days or hours between events. This makes interpreting the chart unusual. For example, if infection rate *increases*, the time or number of intervals between infections would *decrease* and could even be as low as 0. If the rate *decreases*, the time or number of intervals between infections would *increase*. Thus, a point beyond the upper control limit would indicate an unusually long period of time between infections—in other words, that the rate was unusually low.

One negative property of the chart is that, if the control limits are fixed and only Test 1 is used, the Average Run Length (ARL) will increase if the rate increases. If the rate increases by 25%, and the control limits are fixed, the ARL will increase by approximately 40%. Therefore, the T chart will be slow to detect increases in the event rate. To compensate, Minitab uses by default both Test 1 and Test 2. Adding Test 2 increases the ARL by only a very small amount for decreases in the average time of around 10% and decreases the ARL for larger changes in the average time.

Calculations Used

Notation

There are 3 types of data that a T chart can be used for:

1. Numeric, non-negative data

   This type of data is the number of intervals between events. It may be continuous (for example, 13.0957) or integer (although integer data is usually associated with the G chart). If this type of data are entered, they are the X_i values used in the chart.

   Note: 0 is an acceptable value. It implies that 2 events occurred at the exact same time. If there are 0's in the data, we use an alternative method for estimating the parameters.

2. Date/time data (for example, 01/23/2011 8:32:14)

   This type of data records the date and time of each event. Each data value must be >= the preceding value. It is acceptable to have dates only, without the time portion (although date-only data is usually associated with a G chart). If dates/times are entered, the X_i values for the chart are calculated as follows:

   Let D_1, D_2, …, D_N be the date/time values entered. Then $X_2 = D_2 – D_1$, $X_3 = D_3 – D_2$, … , $X_N = D_N – D_{N-1}$.

   The resulting data are the (integer or non-integer) number of days between events.

   Note: If only dates are entered, the resulting days-between data are integers. This type of data is often associated with the G chart.

3. Time-between data (for example, 8:32:14) is also known as elapsed-time data. The data represent the elapsed time between event i and event i-1. If this type of data is entered, they are the X_i values in the chart.

X_i = plot points, as explained above

If there are no 0's in the X_i data, the MLE estimates of the shape (KAPPA) and scale (LAMBDA) parameters are calculated from the data and used to obtain the percentiles of the Weibull distribution.

If there are 0's in the X_i data, the following alternative method for obtaining parameters is used:

1) Rank data lowest to highest.
2) p = (rank – 0.3) / (n + 0.4)
3) X = ln(-ln(1 – p))
4) Remove rows (from both X and Data) where Data = 0
5) Y = Ln(Data)
6) Regress Y on X, obtaining the equation $Y = B_0 + B_1*X$
7) LAMBDA = $\exp(B_0)$, KAPPA = $1/B_1$

Limits estimated from data

Let p1, p2, p3, p4, p5, p6, p7 be the CDF values from a Normal(0,1) for –3, -2, -1, 0, +1, +2, +3.

Let w1, w2, w3, w4, w5, w6, w7 be the invcdf values for p1, p2, p3, p4, p5, p6, p7 using a Weibull (KAPPA, LAMBDA) distribution.

Then, get LCL and UCL as follows:

CL = w4

UCL = w7

LCL = w1

Using historical parameters

If historical parameters are specified, the chart is based on the shape and scale parameters of the Weibull distribution, much like other charts use the mean and standard deviation. One difference is that the user must enter historical values for both parameters (in charts like I Chart of Xbar Chart they can enter one or both parameters).

The shape parameter must be > 0, and in most cases it should be between 0.5 and 2, although these limits are used primarily for practical reasons. Shape parameters < 0.5 imply a distribution that is extremely skewed and can have a kurtosis value that exceeds 2000. (An exponential distribution has a kurtosis value of only 6.) Shape parameters that are higher than 2 imply a distribution that is approaching symmetric, or even left skewed. Both are quite unrealistic because data for the time between events is usually highly skewed to the right.

The scale parameter must be > 0 and should be somewhat greater than the mean of the data. If the scale parameter is less than the mean of the data or too much greater than the mean, the limits on the chart will not reflect the process accurately and could lead to many false alarms.

Note: The historical values entered replace the KAPPA and LAMBDA used in the equations above to obtain the control limits, center line, etc.

Standard Control Chart Tests Used in the T Chart

Test 1 – 1 point outside percentiles corresponding to K standard deviations away from the center line in a chart based on the normal distribution (plot point < w1 or > w7, if K = 3, see below if K <> 3)

Test 2 – K points in a row on one side of the center line

Test 3 – K points in a row, all increasing or decreasing

Test 4 – K points in a row, alternating up and down

Test 5 – K out of K + 1 points > w6, or K out of K + 1 points < w2

Test 6 – K out of K + 1 points > w5, or K out of K + 1 points < w3

Test 7 – K points in a row >= w3 and <= w5

Test 8 – K points in a row < w3 or > w5

For Test 1, if the argument K is 3, then the w1 and w7 values used for the control limits are used to define Test 1 failures (i.e., points that are < w1 or > w7). If the argument K is not equal to 3, then define p1' and p2' as the cdf values of Normal(0,1) for –K and +K. Then define w1' and w7' as the invcdf values from Weibull (KAPPA, LAMBDA) corresponding to p1' and p2'. The definition of a test 1 failure is then a point < w1' or > w7'.

In the tests above, w1, w2, w3, w4, w5, w6, w7 are as defined earlier (i.e., invcdf values from Weibull distribution corresponding to p1, p2, p3, p4, p5, p6, p7 the cdf values of Normal(0,1) for -3, -2, -1, +1, +2, +3. However, if the Test 1 argument is <> 3 we replace only w1 and w7 with w1' and w7'.

References

[1] L. Y. Chan, D. K. J. Lin, M Xie, and T. N. Goh. Cumulative Probability Control Charts for Geometric and Exponential Process Characteristics. *International Journal of Production Research*, 40:133-150, 2002.

[2] L. Y. Chan, M Xie, and T. N. Goh. Cumulative Quantity Control Charts for Monitoring Production Processes. *International Journal of Production Research*, 38:397-408, 2000.

[3] F. F. Gan. Design of Optimal Exponential CUSUM Charts. *Journal of Quality Technology*, 26(2):109-124, 1994.

[4] F. F. Gan. Designs of One-Sided and Two-Sided Exponential EWMA Charts. *Journal of Quality Technology*, 30:55-69, 1998.

[5] F. C. Kaminsky, J. C. Benneyan, R. D. Davis, and R. J. Burke. Statistical Control Charts Based on a Geometric Distribution. *Journal of Quality Technology*, 24:63-69, 1992.

[6] J. Y. Liu, M. Xie, T. N. Goh, and P. R. Sharma. A Comparative Study of Exponential Time Between Events Charts. *Quality Technology of Quantitative Management*, 3:347-359, 2006.

[7] D. C. Montgomery. *Introduction to Statistical Quality Control*, Wiley, 6th Edition, 2009.

[8] M. Xie, T. N. Goh, and P. Ranjan. Some Effective Control Chart Procedures for Reliability Monitoring. *Reliability Engineering and System Safety*, 77:143-150, 2002.

[9] C. W. Zhang, M. Xie, and T. N. Goh. Design of Exponential Control Charts Using a Sequential Sampling Scheme. IIE Transactions, 38:1105-1116, 2006.

[10] C. W. Zhang, M. Xie, and T. N. Goh. Economic Design of Exponential Charts for Time Between Events Monitoring. *International Journal of Production Research*, 43:5019-5032, 2005.

Prepared by Dr. Terry Ziemer, SIXSIGMA Intelligence