

Potpourri ...

This session (only the first part of which is covered on Saturday AM ... the rest of it and Session 6 are covered Saturday PM) is an amalgam of several topics. These are

1. confidence limits for σ , the standard deviation of a normal population
2. some qualitative discussion of elementary statistics and business process improvement efforts and programs
3. Shewhart "control charts" for means and standard deviations

Inference for σ

One additional method of one-sample statistical inference is that for σ . MMD&S don't cover this method because it is thought to be non-robust (the nominal confidence level doesn't extend to inference for non-normal populations). While that may be true, having this method available is better than having no method. Besides, practice with it gives a sense for how little (relatively speaking) one learns about population standard deviations from small samples (as compared to how much one learns about means).

The basic probability fact that enables inference for a population standard deviation is

When sampling from a normal population/universe/process, the random quantity

$$\frac{(n - 1) s^2}{\sigma^2}$$

has a famous standard probability distribution called the " χ^2 distribution with $df = n - 1$ "

The χ^2 distributions are right-skewed distributions over positive values (obviously, the quantity above can never be negative!). Plots of several χ^2 densities are below.

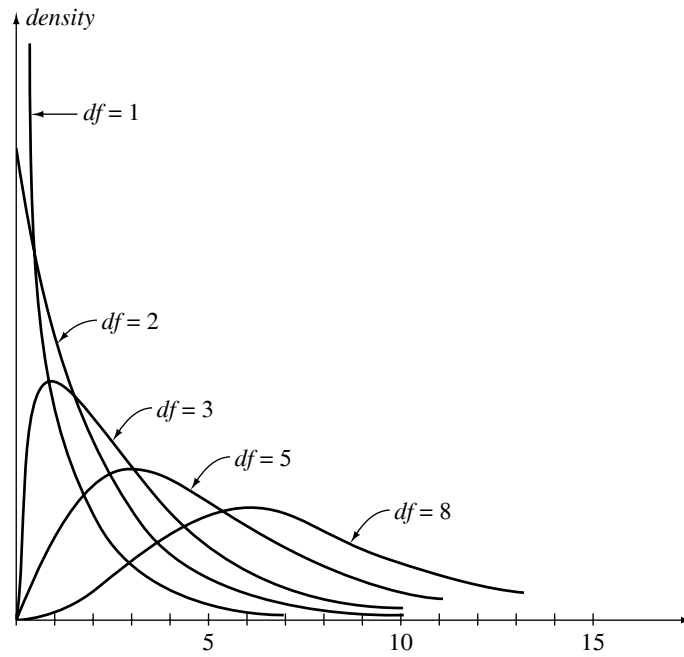
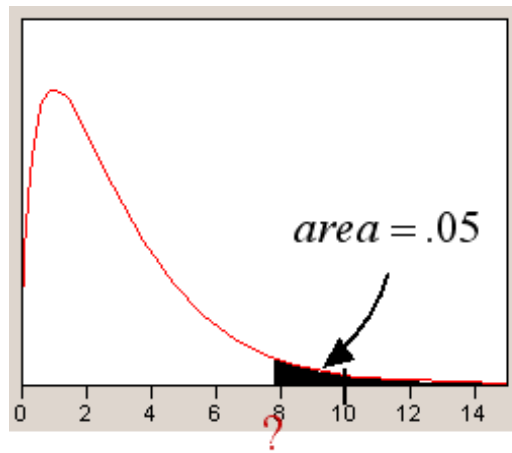


Figure 1: Several χ^2 Densities

MMD&S have a χ^2 table that isn't quite adequate for our present purposes. A more complete one (set up in the same fashion as MMD&S Table F on page T-20) is on the Stat 328 Web site on the Handout page. It and the MMD&S table are set up like the t table. One locates a desired *right tail area* across the top margin of the table, a desired degrees of freedom down the left margin of the table, and reads out a desired cut-off value from the body of the table. This is illustrated in schematic fashion below.



χ^2 distribution with $df = 3$

df	.05
3	7.81

Figure 2: Layout of χ^2 Table

Example Consider how sample standard deviations, s , based on samples of size $n = 5$ from the brown bag (Normal with $\mu = 5$ and $\sigma = 1.715$) will vary. The new probability fact says that

$$\frac{(5 - 1) s^2}{(1.715)^2} = 1.36s^2$$

has the χ^2 distribution with $df = 5 - 1 = 4$. So, for example, from the χ^2 table we can see that

$$P\left(.711 \leq 1.36s^2 \leq 9.488\right) = .90$$

(the lower 5% point of the distribution is .711 and the upper 5% point is 9.488).

I can manipulate the new probability fact to get confidence limits for σ . To do this, I first choose numbers L, U so that

$$P\left(L < \frac{(n - 1) s^2}{\sigma^2} < U\right) = \textit{desired confidence level}$$

(For example, if I'm using 90% confidence and $n = 5$, $L = .711$ and $U = 9.488$ do the job.) Then having

$$L < \frac{(n-1)s^2}{\sigma^2} < U$$

is equivalent to having

$$\frac{(n-1)s^2}{U} < \sigma^2 < \frac{(n-1)s^2}{L}$$

which makes

$$\frac{(n-1)s^2}{U} \quad \text{and} \quad \frac{(n-1)s^2}{L}$$

confidence limits for σ^2 . Then taking square roots, one has the confidence interval for σ

$$\left(\sqrt{\frac{(n-1)s^2}{U}}, \sqrt{\frac{(n-1)s^2}{L}} \right)$$

i.e.

$$\left(s\sqrt{\frac{(n-1)}{U}}, s\sqrt{\frac{(n-1)}{L}} \right)$$

Example Consider once more the brown bag, and this time estimation of the standard deviation based on a sample of size $n = 5$. A 90% confidence interval for σ (known in this artificial context to be 1.715) is

$$\left(s\sqrt{\frac{5-1}{9.488}}, s\sqrt{\frac{5-1}{.711}} \right)$$

that is,

$$(.65s, 2.37s)$$

Example Consider once more the 1999 no-load mutual fund rate-of-return data of Dielman. Let's make (again assuming that the Dielman data is a random

sample from a relatively bell-shaped population of such rates) a 95% confidence interval for σ , the variability in the rates. Since $n = 27$ for this problem, we enter the χ^2 table using $df = 26$. Recalling that $s = 4.1$, the confidence interval is then

$$\left(s\sqrt{\frac{(n-1)}{U}}, s\sqrt{\frac{(n-1)}{L}} \right)$$

that is,

$$\left(4.1\sqrt{\frac{26}{41.923}}, 4.1\sqrt{\frac{26}{13.844}} \right)$$

or

$$(3.2, 5.6)$$

Exercise Consider again the scenario of Problem 7.31, page 456 of MMD&S. For that scenario, make a 90% confidence interval for the population standard

deviation of corn prices. Remember that in that problem $n = 22$ and

$$.176 = \frac{s}{\sqrt{n}} = \frac{s}{\sqrt{22}}$$

so that

$$s = .176\sqrt{22} = .83$$

A Bit About Basic Statistics and Business Process Improvement (Some Context for the Material of Ch 12 MMD&S)

The obvious fact of global competition has made ever-increasing efficiency/effectiveness a major concern of businesses everywhere. Some main features of what has evolved over the past 2 or 3 decades in this regard are:

- modern business improvement strategies and programs are "process-oriented"
- "continuous improvement" is standard (essentially self-evident) doctrine
- process improvement strategies are necessarily data-driven *and therefore intrinsically statistical*
- there are many flavors/lists of steps/sets of jargon in this area ... good ones must boil down to sensible, logical, data-driven problem solving ... the application of the "scientific method" to business

Popular flavors of improvement programs have included

- Six Sigma (Motorola/GE etc.)
- ISO 9000 (specifications and standard language for quality systems)
- Malcolm Baldrige ("Baldrige process" in pursuit of a "Baldrige prize" ... the US version of the Deming Prize)
- TQM
- Deming's "14 points"
- "Theory of Constraints"

- "Zero Defects"
- etc./etc./etc.

There are by now thousands of articles and books that have been written on these flavors ... a Web search on any of these phrases will produce hundreds of leads, many to consultants who will gladly sell "their own" "proprietary" versions of programs in this area.

Of the above flavors of business process improvement movement, "Six Sigma" is probably currently the most fashionable. The name "Six Sigma" refers to at least 3 different things:

1. a goal for process performance (that essentially guarantees that all process outcomes are acceptable)
2. a discipline or methodology for improvement aimed at achieving the goal in 1)
3. an organization/program for training in and deployment of the discipline in 2)

The Six Sigma goal for process performance is that:

The mean process outcome is at least 6σ below the maximum acceptable outcome and at least 6σ above the minimum acceptable outcome.

This can be pictured as:

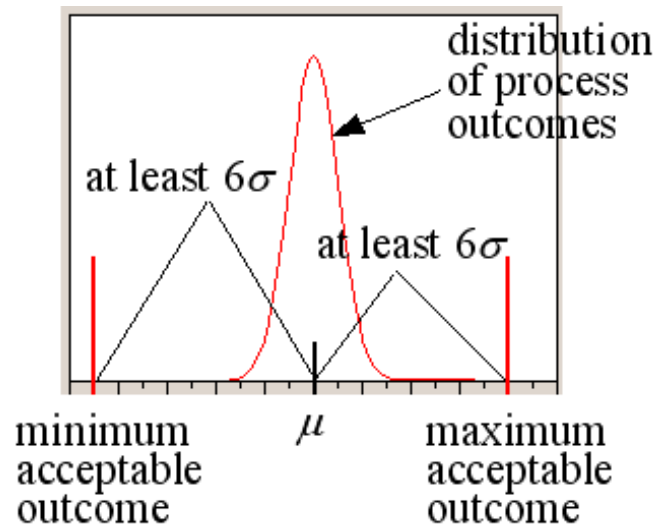


Figure 3: The Six Sigma Goal for Process Performance

The Six Sigma problem solving/process improvement paradigm is organized around the acronym

(D)MAIC

(Define)

Measure

Analyze

Improve

Control

- The "Define" step of the paradigm involves agreeing upon the scope of an improvement project. What are the boundaries?
- The "Measure" step includes identifying appropriate indicators of process performance and potentially influential drivers of that performance, developing measurement systems, and collecting initial data.

- The "Analyze" phase includes doing an initial analysis of process data (both descriptive and inferential) and collecting additional data as needed. The goal here is to get an initial picture of process performance.
- The "Improve" step involves bringing to bear whatever tools of logic, experimentation, technology, collaboration, organization, etc. that seems most likely to help.
- The final "Control" step of the paradigm involves setting up a system to monitor newly improved process performance into the future. The intent is to verify that the improvement persists, that problems fixed "stay fixed."

The Six Sigma system of deployment of the (D)MAIC paradigm involves training (a fair amount in the kind of basic statistics just covered in Sessions 1-4),

organization into project teams, and recognition of training and project success. (Project success is always supposed to be measured in dollar impact.) For whatever reason, martial arts terminology has been used in Six Sigma programs to indicate achievement levels. Neophytes in the system first earn their "green belts." Experienced and effective individuals become "black belts" and "master black belts."

Shewhart "Control" Charts ... Statistical Process Monitoring

The final step in the Six Sigma (D)MAIC paradigm is that of "control" in the sense of monitoring for purposes of detecting process change. Chapter

12 of MMD&S principally concerns standard tools for this activity, so called "Shewhart control charts." These date to Walter Shewhart's work in the late 1920's at Bell Labs.

The basic motivation/idea of Shewhart charting is that one wants consistency of process behavior over time ... but recognizes that complete consistency is impossible, too much to hope for in the real world. However, it is not too much to hope for that a process has an pattern of variation that is unchanging over time. That is, one might hope that process data look like random draws from a fixed universe/population. Shewhart charts are tools meant to monitor for departures from that kind of "sampling from a fixed distribution" behavior.

The simplest version of a Shewhart control chart is the Shewhart \bar{x} chart. These are set up like

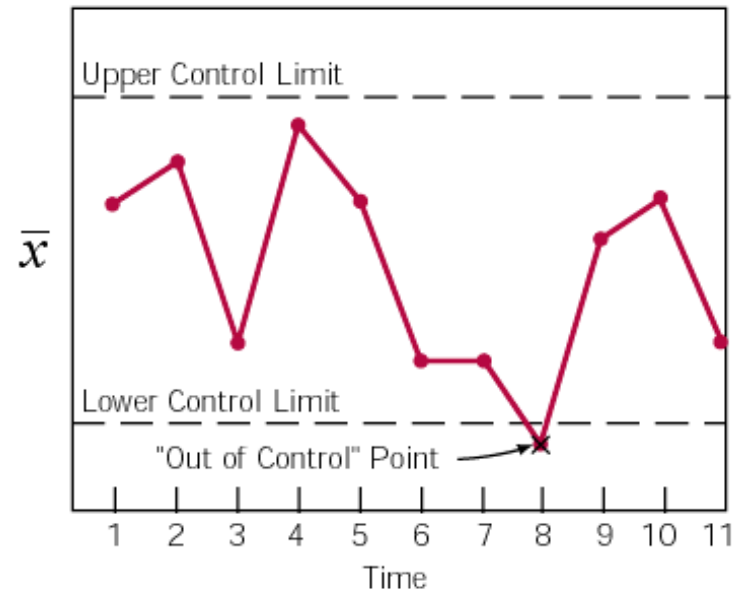


Figure 4: A Shewhart \bar{x} Chart

"Control Limits" separate values of \bar{x} that are plausible if indeed process output is acting like random draws from a fixed universe from those that are implausible under that mental model. Where could such limits come from? What we know about the probability distribution of \bar{x} can be used. If standard/expected process behavior can be described by a mean μ and a standard deviation σ , material from Session 2 tells us that \bar{x} has a normal distribution with

$$\mu_{\bar{x}} = \mu \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Then the 68%/95%/99.7% rule says that values of \bar{x} between

$$\mu_{\bar{x}} - 3\sigma_{\bar{x}} = \mu - 3\frac{\sigma}{\sqrt{n}} \text{ and } \mu_{\bar{x}} + 3\sigma_{\bar{x}} = \mu + 3\frac{\sigma}{\sqrt{n}}$$

are in some sense "plausible" under a "no process change" model, while others are not. That suggests using

$$LCL_{\bar{x}} = \mu - 3\frac{\sigma}{\sqrt{n}} \text{ and } UCL_{\bar{x}} = \mu + 3\frac{\sigma}{\sqrt{n}}$$

This can be pictured as

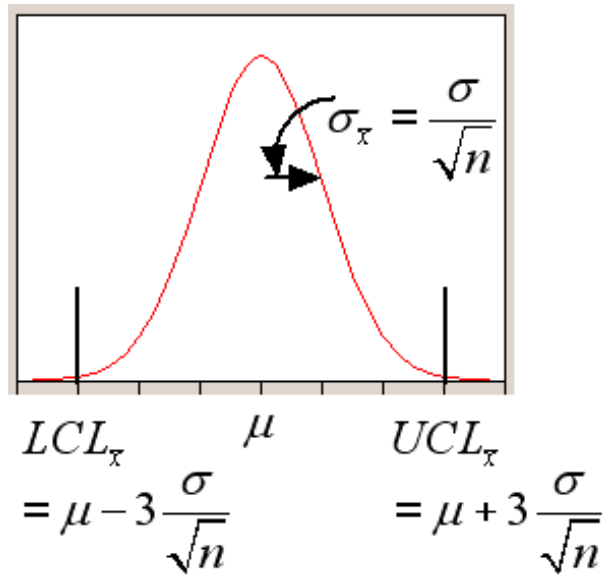


Figure 5: Control Limits for a Shewhart \bar{x} Chart

Example A process improvement team in a very large office invented a 0-100 scale for scoring the correctness with which a type of company form was filled out and filed. After a training period for the large number of workers responsible for filling out these forms, a large number of these forms were audited for correctness and scored using the 0-100 scale (100 was best). The scores of the audited forms had a left-skewed distribution with mean 93.44 and standard deviation 17.96. A JMP report is below.

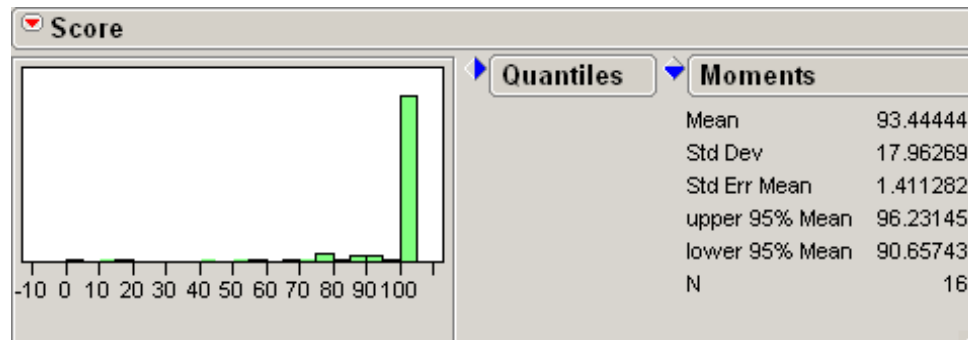


Figure 6: JMP Report of Scores in a Large Paperwork Correctness Audit

Suppose that henceforth, monthly samples of $n = 50$ of these forms will be selected at random (from the very large number handled in the office) and the sample mean score computed. What might then be appropriate control limits for \bar{x} ? If the pattern of variation seen immediately after the training period continues into the future, we may approximate the probability distribution for an average of $n = 50$ scores as normal, with

$$\mu_{\bar{x}} = \mu \approx 93.44 \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{17.96}{\sqrt{50}} = 2.54$$

A lower control limit for future sample means (based on $n = 50$) is therefore

$$LCL_{\bar{x}} = 93.44 - 3(2.54) = 85.82$$

and since $93.44 + 3(2.54) = 101.06$ while no average can exceed 100, there will be no upper control limit in this application. (Note that larger sample sizes WOULD produce an upper *control limit* for \bar{x} less than 100.)

Notice that

1. control limits apply to plotted statistics (like \bar{x}), not to individual values from a process
2. control limits say absolutely nothing about the acceptability or functionality of individual values
3. control limits simply tell one if a process has changed

Regarding 2) and 3), note that in the paperwork auditing example, for sample sizes larger than $n = 50$, an unexpected process *improvement* (seen in a large \bar{x}) can show up as an "out-of-control" \bar{x} .

AAO Unit 18 (47:03) \bar{x} Charts at Frito-Lay/Deming

Exercise Problem 12.6 page 12-14 of MMD&S. (Note that here $n = 4$, the number of items checked at a given period. $\mu = 75$, the ideal value, and $\sigma = .5$ is available from past experience.

There is a second type of Shewhart control chart introduced in Sections 12.1 and 12.2 of MMD&S. This is the " s chart," made by plotting sample standard deviations. (One uses such a chart to check the constancy of process short term variation.) In order to find appropriate control limits for s one needs the kind of information about the distribution of s that led to our confidence limits for σ . That kind of information leads to the following facts.

When sampling from a Normal universe, both the mean and standard deviation of s are proportional to the population standard deviation,

σ . The constants of proportionality are called c_4 and c_5 (and can be found on page 12-19 of MMD&S). That is

$$\mu_s = c_4\sigma$$

$$\sigma_s = c_5\sigma$$

(Don't get lost in the notation here. This just says that the average sample standard deviation grows in proportion to the variability of the population being sampled and that simultaneously, the variability of the sample standard deviation grows in proportion to the population variability.)

Then using the common Shewhart control chart "3 sigma limits" idea, common

control limits for s become

$$\begin{aligned}UCL_s &= \mu_s + 3\sigma_s \\&= c_4\sigma + 3c_5\sigma \\&= (c_4 + 3c_5)\sigma \\&= B_6\sigma\end{aligned}$$

and

$$\begin{aligned}LCL_s &= \mu_s - 3\sigma_s \\&= c_4\sigma - 3c_5\sigma \\&= (c_4 - 3c_5)\sigma \\&= B_5\sigma\end{aligned}$$

where B_5 and B_6 are tabled on page 12-19 of MMD&S.

Example Consider setting up control charts to monitor sampling from a process conceptually equivalent to the brown bag (i.e. a normal process with mean $\mu = 5$ and $\sigma = 1.715$) based on sample of size $n = 5$.

An \bar{x} chart would have a center line at

$$\mu = 5$$

an upper control limit at

$$\begin{aligned} UCL_{\bar{x}} &= \mu + 3\frac{\sigma}{\sqrt{n}} \\ &= 5 + 3\frac{1.715}{\sqrt{5}} \\ &= 7.3 \end{aligned}$$

and a lower control limit at

$$\begin{aligned} LCL_{\bar{x}} &= \mu - 3\frac{\sigma}{\sqrt{n}} \\ &= 5 - 3\frac{1.715}{\sqrt{5}} \\ &= 2.7 \end{aligned}$$

A corresponding s chart would have a center line at

$$\mu_s = c_4\sigma = (.9400) 1.715 = 1.612$$

an upper control limit at

$$UCL_s = B_6\sigma = (1.964) 1.175 = 3.368$$

and since B_5 for sample size turns out to be negative, there is no lower control limit for s for this sample size. (s is never negative.)

A second version of the \bar{x}/s control charting business is what is commonly known as "retrospective" or "as past data" control charting. See Section 12.2 of MMD&S. The idea here is that one is *not* operating in real time with known values μ and σ giving limits to apply to \bar{x} and s as data come in. Rather, "after the fact" one is given a set of data and asked

Do these data look like they could have come from a physically stable process/a single fixed universe?

To answer this question people

1. make a provisional assumption that the process *was* stable/unchanged over the period of data collection
2. on the basis of 1) use the data in hand to estimate μ and σ
3. plug the estimates from 2) into formulas for control limits for \bar{x} and s and use them to criticize 1)

A reasonable question is exactly what one should use for the estimates in 2). To approximate μ it is standard to use

$$\hat{\mu} = \bar{\bar{x}}$$

(the average sample mean) and to approximate σ , it is standard to use

$$\hat{\sigma} = \frac{\bar{s}}{c_4}$$

(the average sample standard deviation divided by the constant c_4). Plugging these estimates into the formulas on page 12-18 of MMD&S, one gets the "retrospective" formulas on page 12-32.

JMP will produce both the (practically far more useful) control charts where μ and σ are furnished (and in reality applied in real time as samples roll in) and also the retrospective charts. The real time control charts can be used to ask:

"Does it seem that the process was stable during this present period at the given values μ and σ ?"

The retrospective charts can be used to ask:

"Does it seem that the process was stable at *some* values μ and σ *over the whole length of data collection*?"

Again, the point of Shewhart control charting is "process monitoring for purposes of change detection."