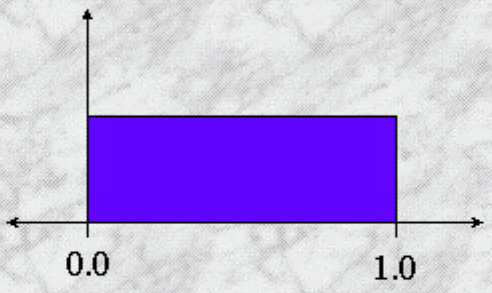
Revisions to the CLIGEN Weather Generator

Charles R. Meyer
USDA-ARS-NSERL
West Lafayette, IN

Topics Addressed:

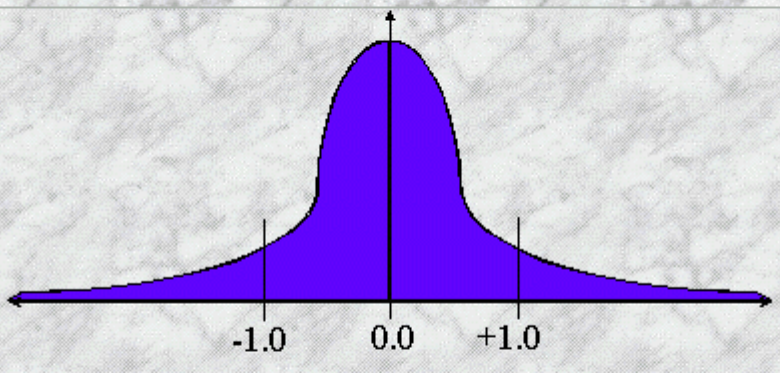
- **Some essential concepts.**
- How CLIGEN works.
- What was going wrong.
- What's been done about it.
- What effect the changes have.

Uniform Random Number Generator:



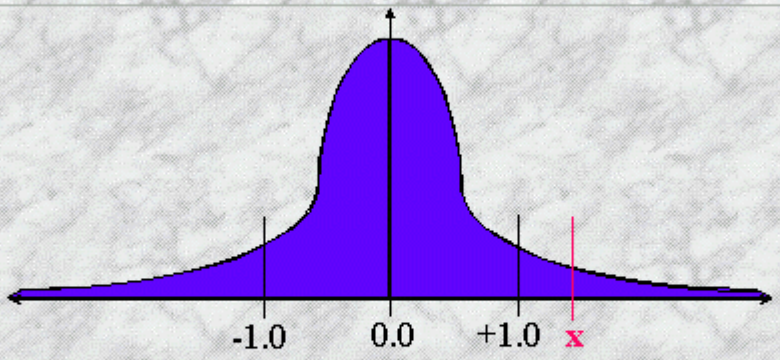
Creates a distribution of numbers, in which values between 0 and 1 are equally likely to occur.

Standard Normal Deviate Generator:



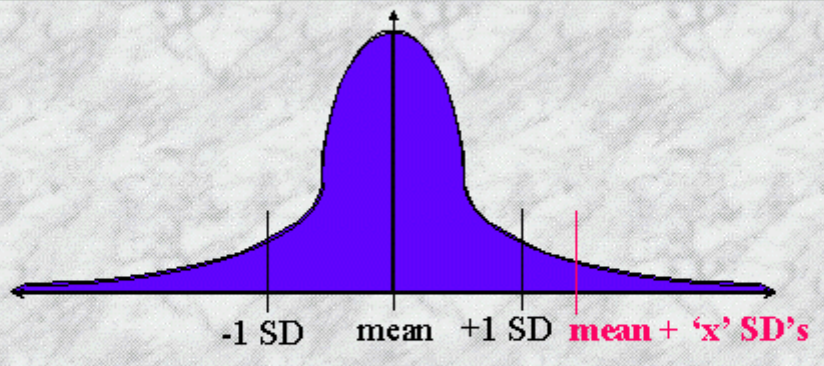
Uses random numbers to generate a distribution that approximates standard normal (mean = 0, SD = 1).

Converting from Standard Normal . . .



To convert a point '**x**' SD's from the mean on the standard normal curve, to the equivalent point on a generic normal curve . . .

. . . to "generic" Normal:



. . . multiply the number of SD's from the mean (positive or negative), times the generic SD, and add the result to the generic mean.

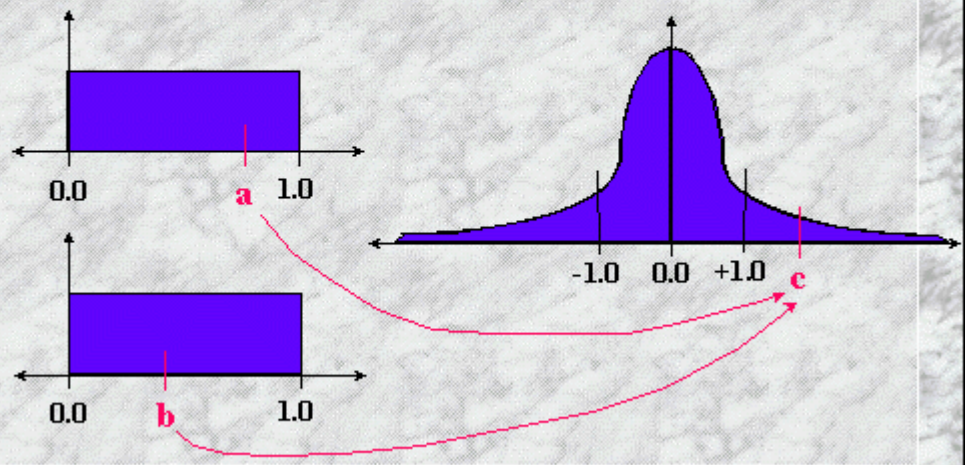
Topics Addressed:

- Some essential concepts.
- **How CLIGEN works.**
- What was going wrong.
- What's been done about it.
- What effect the changes have.

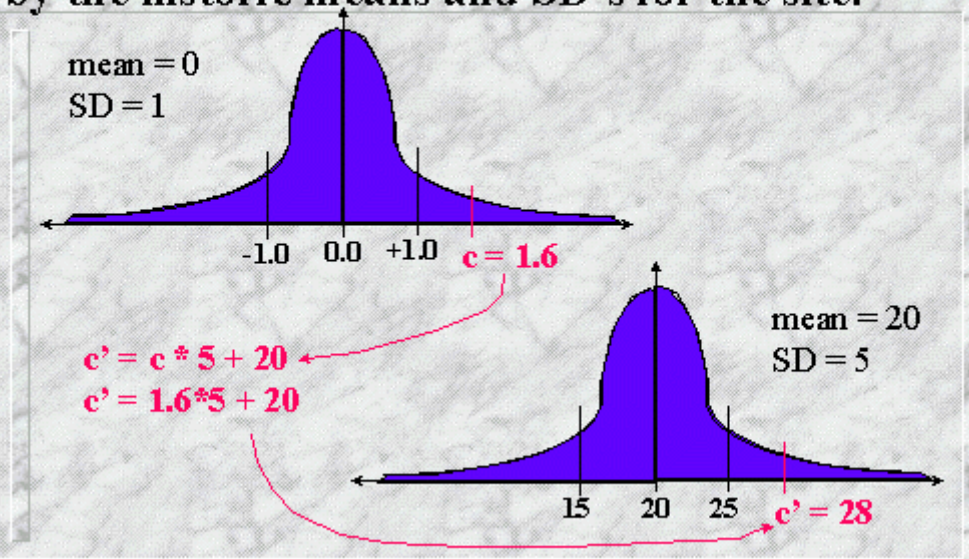
CLIGEN uses a separate copy of its random number generator for each of 9 parameters:

- Maximum Temperature
- Minimum Temperature
- Dewpoint Temperature
- Radiation
- Probability of Precipitation
- Amount of Precipitation
- Time to Peak Intensity
- Wind Velocity
- Wind Direction

The random numbers are fed pair-wise into a "standard normal deviate generator" to produce daily standard normal values:

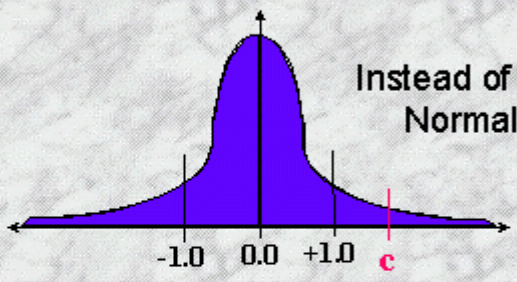


For temperatures and dewpoints the daily standard normal values generated are "scaled" by the historic means and SD's for the site:

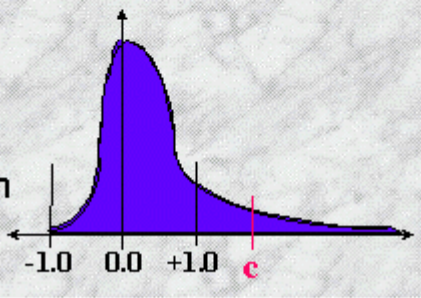


Precip data is a bit more complicated . . .

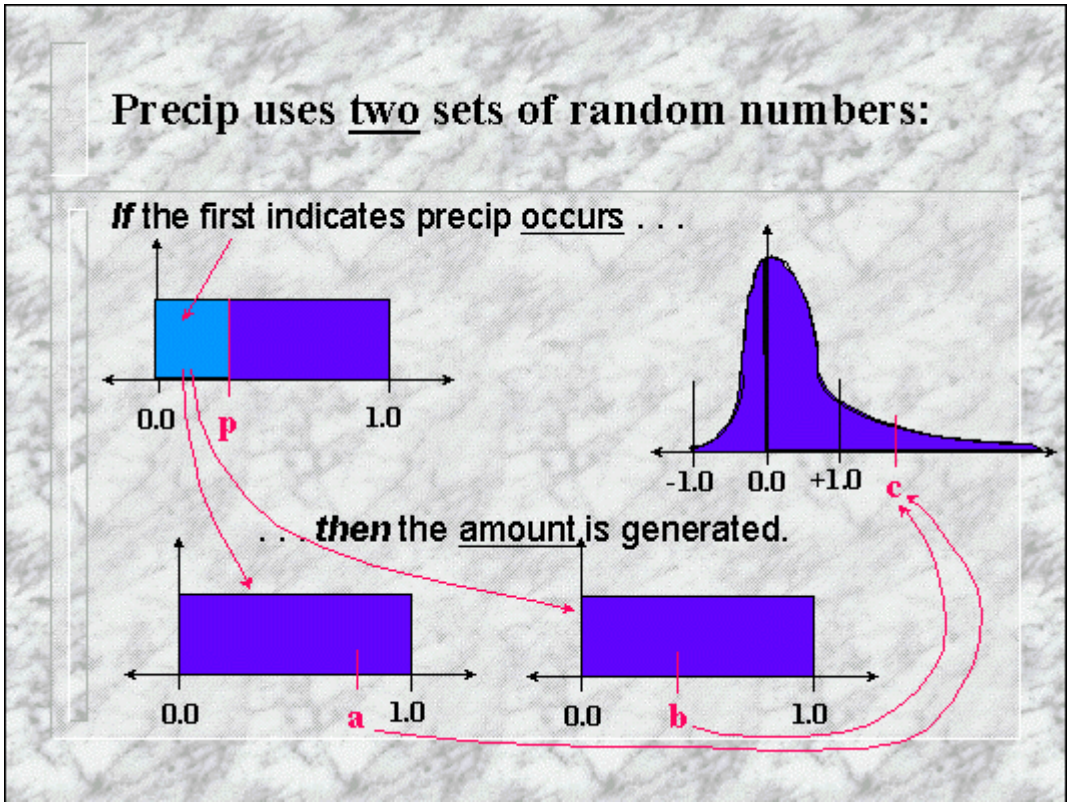
Instead of following a simple Normal distribution . . .



. . . precip follows a "skewed" distribution, and a Pearson Type III equation is used.



Precip uses two sets of random numbers:



Topics Addressed:

- Some essential concepts.
- How CLIGEN works.
- **What was going wrong.**
- What's been done about it.
- What effect the changes have.

The Problem:

- CLIGEN uses monthly means, standard deviations, and skewness, of historical data collected at the site.
- Using the Central Limit Theorem from statistics, it is possible to perform interval tests on the monthly means generated, to determine how well they match the historical means from which they came.
- More than the expected percent of the time, CLIGEN's outputs were not matching its inputs.

More on the Problem:

- CLIGEN's outputs are derived using a uniform random number generator (RNG) and a standard normal deviate generator (SNG) intended to produce random sets of numbers approximating a standard normal distribution.
- It is possible to perform tests on both the mean and SD of a set of numbers to determine how well they match a standard normal distribution.
- The tests showed the numbers often did a very poor job of approximating a standard normal distribution.

Topics Addressed:

- Some essential concepts.
- How CLIGEN works.
- What was going wrong.
- **What's been done about it.**
- What effect the changes have.

What was done:

- If the RNG and SNG are perceived as a "factory" producing numbers for CLIGEN to consume, one would say some of their units don't meet production standards.
- In a physical factory, a common approach to quality control is to inspect the items coming off the production line, and reject those not passing inspection.
- An analogous approach was taken with the RNG and SNG in CLIGEN. Tests are imposed on both the monthly mean and the monthly SD of the numbers they produce. Sets not passing both tests are rejected and a replacement is generated.

Topics Addressed:

- Some essential concepts.
- How CLIGEN works.
- What was going wrong.
- What's been done about it.
- **What effect the changes have.**

Results and Implications:

- The outputs of CLIGEN now do a much better (and fully acceptable) job of reproducing means (and SD's) of the data from which they were generated.
- The user can specify a target level of quality. As with a physical factory, the higher the standards, the higher the rejection rate.
- Any tests previously performed to determine how well CLIGEN reproduces extreme events, are not applicable to this version.
- Whether the new method causes a time bias in the numbers generated, is not known. Sets deemed unacceptable early in the run might be accepted later.