ITC Tutorial: Design of Experiments and Analysis of Data

by

Joel Tellinghuisen
Department of Chemistry
Vanderbilt University
Nashville, TN 37235

The Plan of Attack

"5-minute statistics":	Means, variances, and probability distributions
Linear least squares:	The basics
Nonlinear least squares:	What's new?
Application to ITC:	Optimizing parameters by least squares in "experiment design" mode
Systematic Errors:	Is the fit model right?
Beyond the "black box":	Changing things Reporting results

5-Minute Statistics

Sampling | Theory

The mean

$$\bar{x} = \langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad\qquad \langle x \rangle \equiv \mu = \int_{x_{\min}}^{x_{\max}} xP(x)dx$$

The variance

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\delta_i^2 \quad (\delta_i = x_i - \bar{x}) \qquad\qquad \sigma_x^2 = \int_{x_{\min}}^{x_{\max}} (x-\mu)^2 P(x)dx$$

The standard deviation

$$s_x \qquad\qquad \sigma_x$$

The standard error (standard deviation in the mean)

$$\frac{s_x}{\sqrt{n}} \qquad\qquad \frac{\sigma_x}{\sqrt{n}}$$

Probability Distributions

Uniform: $P(x)$ = constant ($a \leq x \leq b$); 0 otherwise

Normal:

$$P_G(\mu,\sigma;x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Poisson: governs counting — $\sigma^2 = \mu$ (= # counts).

Chi-square (χ^2): sampling estimates of variances.

***t*-distribution:** confidence limits for sampling estimates of parameters.

NOTE: Poisson, χ^2, and t-distributions all become Gaussian in the limit of large μ (Poisson) or ν (degrees of freedom, $n-p$).

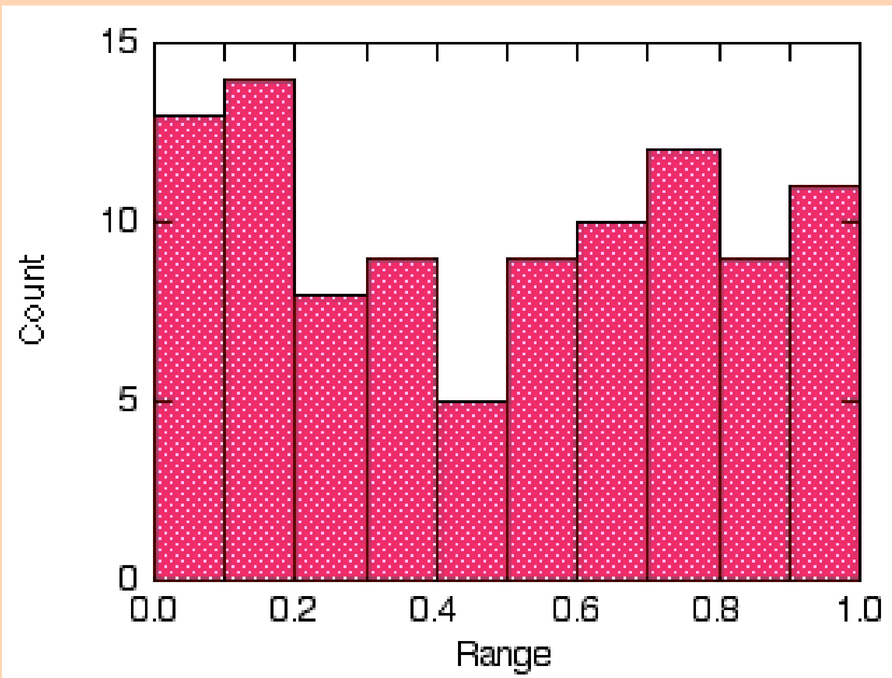
Important probabilities

$$\int_{\mu-\sigma}^{\mu+\sigma} P_G(\mu,\sigma;x)\,dx = 0.683$$

$$\int_{\mu-2\sigma}^{\mu+2\sigma} P_G(\mu,\sigma;x)\,dx = 0.954$$

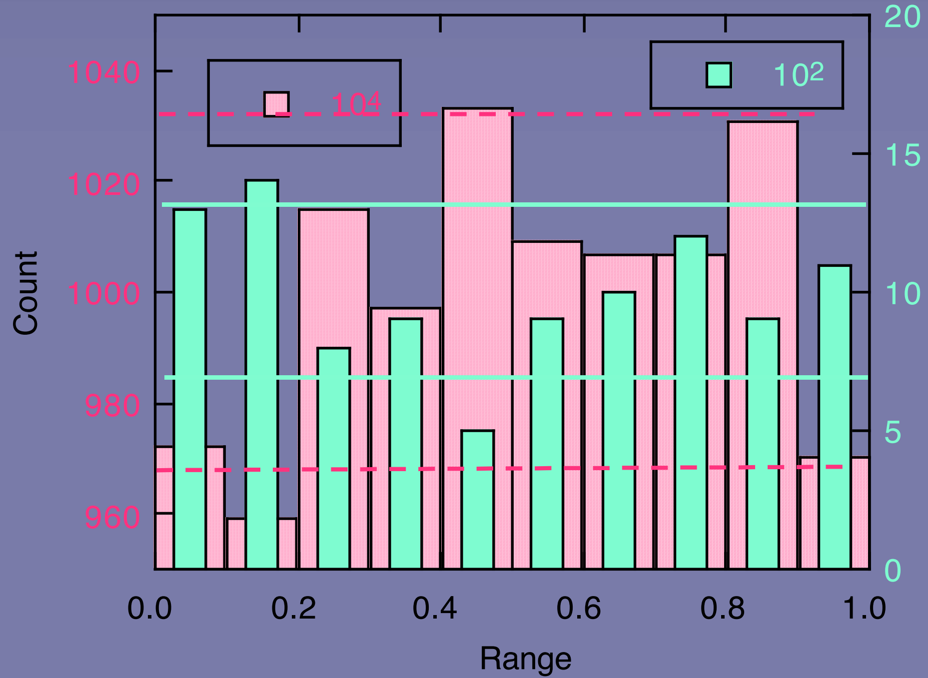
Illustrations

- The uniform distribution is the basis of computer random number generators. By default, the range is $0 < x < 1$, for which $\mu = {}^{1}/_{2}$ and $\sigma^2 = {}^{1}/_{12}$. Let's check …



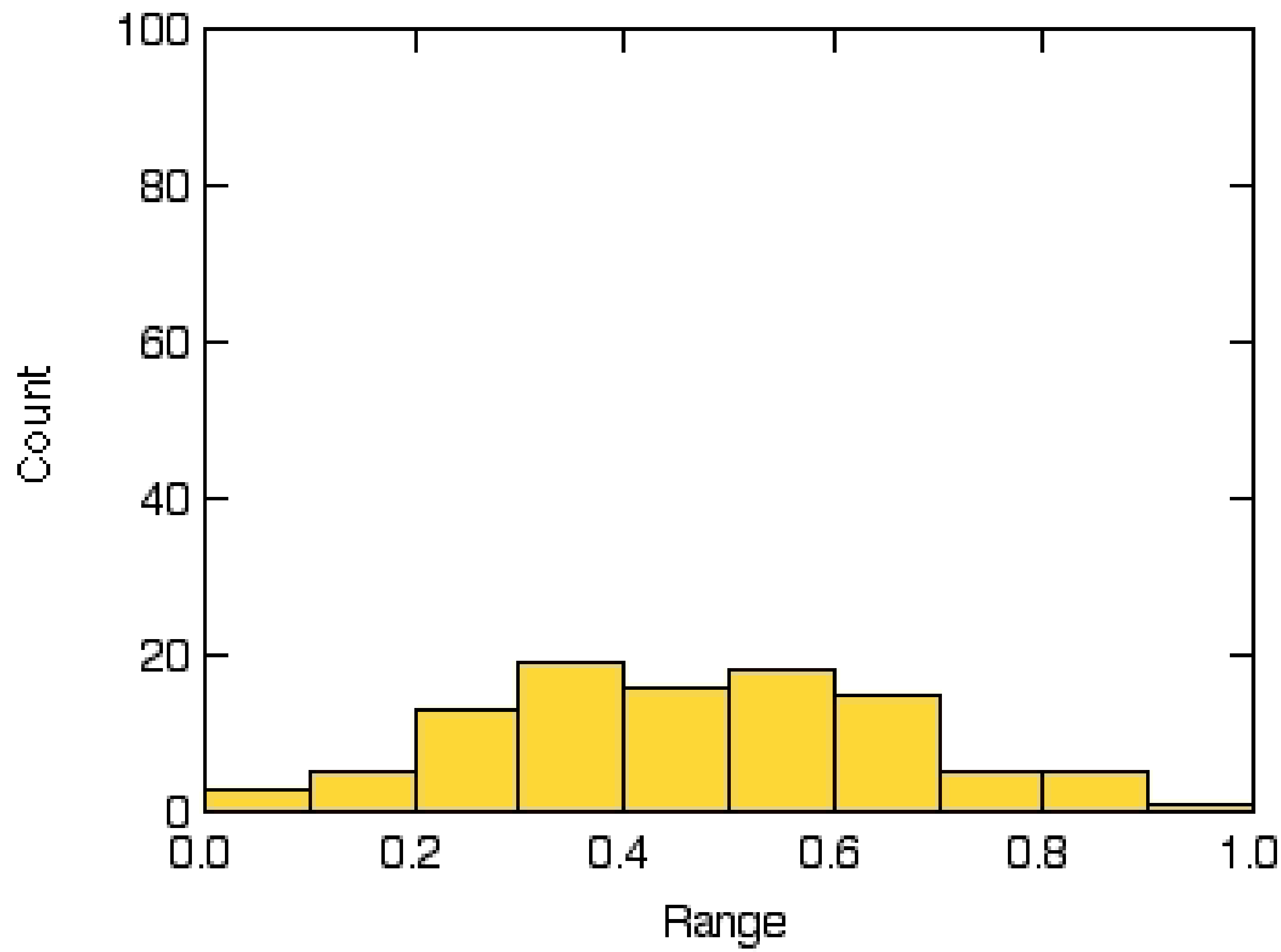
Minimum	9.36e-05
Maximum	0.99998742
Sum	5022.7734
Points	10000
Mean	0.50227734
Median	0.50219405
RMS	0.57845803
Std Deviation	0.28694843
Variance	0.082339402
Std Error	0.0028694843
Skewness	-0.0081906446
Kurtosis	-1.1866255

Closer look at the binning statistics

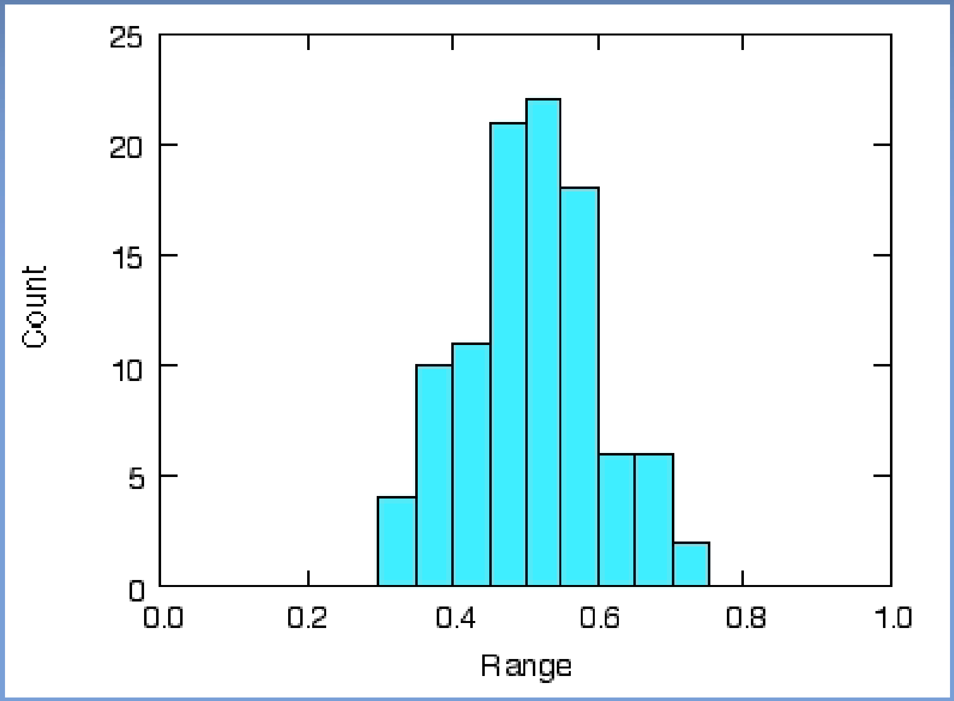


- Compare results for $N = 10^4$ and 10^2.
- Recall □ ≈ n (Poisson).
- Thus about $^2/_3$ of bin counts should fall within ±□ (32,3.2) of the expected values (Gaussian approx.).

Next, bin average of 2 random numbers



And now 12 …



Stats for sum of 12

Minimum	2.6777
Maximum	9.3540
Sum	60038
Points	10000
Mean	6.0038
Median	5.9946
RMS	6.0863
Std Deviation	0.9989
Variance	0.9977
Std Error	0.009988

Results ⇒ the very important **Central Limit Theorem**: Distributions of sums become normal, no matter what the parent distribution, as long as it has finite variance.