

*FinM 331/Stat 339 Financial Data Analysis,
Winter 2010*

Floyd B. Hanson, Visiting Professor

Email: fhanson@uchicago.edu

**Master of Science in Financial Mathematics Program
University of Chicago**

Lecture 1

6:30-9:30 pm, 04 January 2009, Ryerson 251 in Chicago

7:30-10:30 pm, 04 January 2009 at UBS in Stamford

7:30-10:30 am, 05 January 2009 at Spring in Singapore

0.01 Histogram of S&P500 Log>Returns 1988-2008:

Fig04: LR Histogram Vs. NormPDF+1, '88-'08

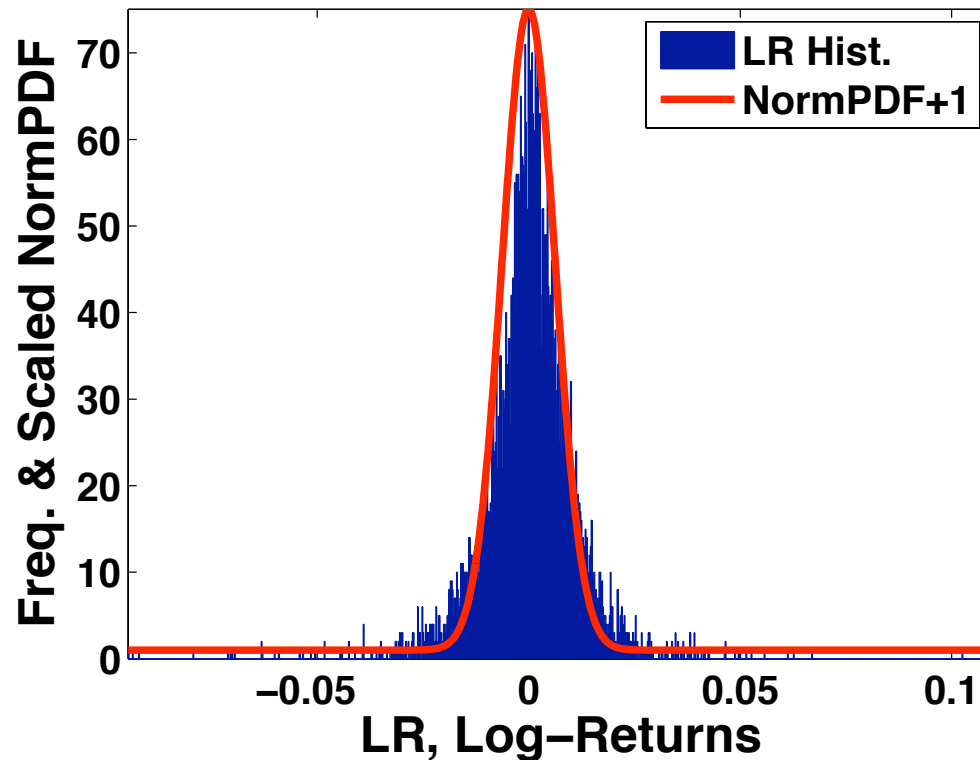
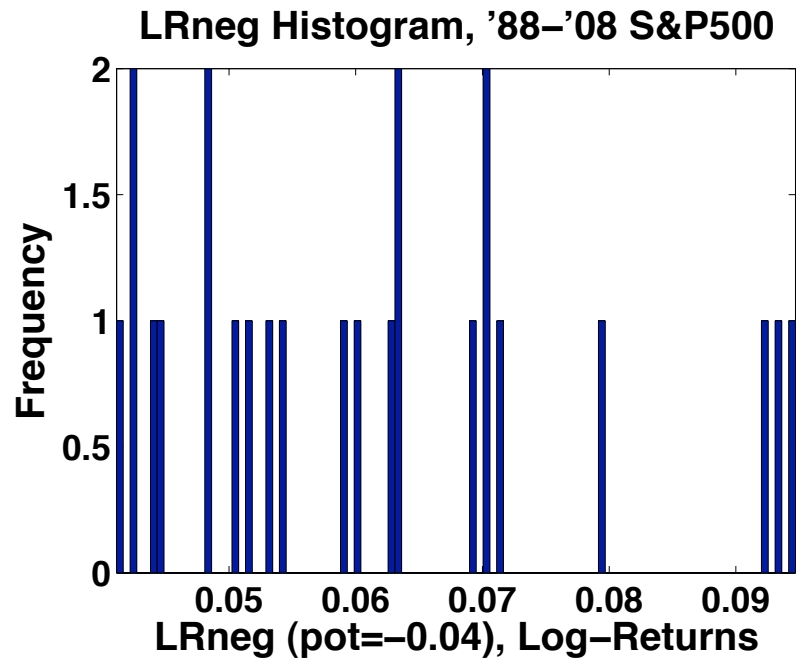
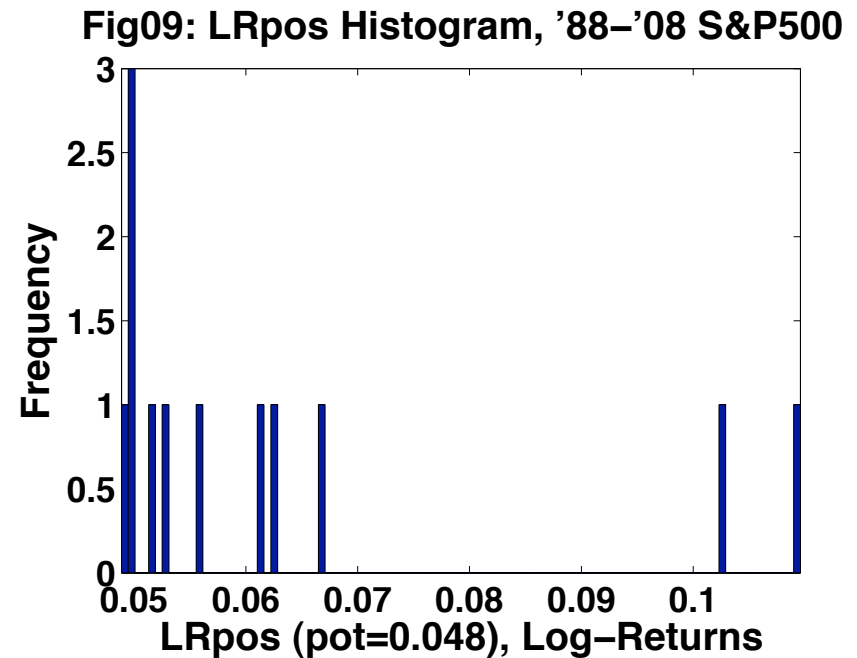


Figure 0.01: S&P500 Daily Log-Return Adjusted Closings from 1988 to 2008 (post-1987) showing long-tails of rare events. Normal kernel-smoothed graph, in **red**, plus one which accounts for non-central and *non-normally invisible, but financially important, rare jumps*.

0.02 Extreme Negative Tail Events for Log-Returns ('88-'08):



(a) Extreme Negative Tails.



(b) Extreme Positive Tails.

Figure 0.02: Extreme Negative and Positive Log-Return Tail Events, with Thresholds $POT = -0.04$ and $+0.048$, respectively. POT means Peaks Over (or Under) Threshold. These represent the *significant crashes or bonanzas* during the time period. {Note: vertical scale differences.}

“... the futility of trying to deal with crises and recessions by finding central bankers and regulators who can identify and puncture bubbles. If these people exist, we will not be able to afford them.”

Robert E. Lucas, Economics, The University of Chicago

In Defense of the Dismal Science, The Economist, Aug 6th 2009.

{Pine-green colored text denotes an active web link.}

Course Outline (tentative)

- 1. Introduction (1 Lecture (L)):** Data models and Discretized Diffusion log-returns, Independent observations and their statistics, Additive and Multiplicative models, Law of large numbers and Monte Carlo applications, Central limit theorem and Extreme tail errors.
- 2. Exploratory Data Analysis (2.5L):** Pseudo-Random Number Generators (RNGs: Uniform, Normal, Exponential, Poisson, Compound Poisson), Histograms, Cumulative histograms, Kernel smoothing, Quantile-Quantile (Q-Q) plots, Discretized Jump-diffusions, Confidence Intervals and Value at Risk (VaR), more Normal distribution poor at extremes, short-fall statistics, Profit and Loss (P&L) statements, Cauchy flat tail distribution, Data estimation, Order statistics, Extreme statistics (Fat tails, Pareto distribution, Peaks Over Thresholds (POTs)).
- 3. Multivariate Statistics (1.5L):** Bivariate distribution, Bivariate Kernel smoothing, Conditional expectation, Covariance, Correlation

Coefficient, Bivariate 3D-histograms, Hybrid fat and thin tail models, Principal Component Analysis (PCA), Multivariate sample means.

- 4. Parametric Regression (1.5L):** Ordinary Least Squares (OLS), Multiple linear regression, MATLAB regression functions, Maximum Likelihood Estimation (MLE), MLE for discretized linear diffusions, compound Poisson processes, MLE for discretized linear jump-diffusions.
- 5. Non-Parametric Regression (1L):** Numerical optimization with derivative-free MATLAB `fminsearch`, Black-Scholes European options, Market calibration and Implied Volatility (IV), Risk-Neutral options and IV, more Kernel smoothing regression.
- 6. General Parameter Estimation (1.5L):** Maximum Likelihood, Moment Methods and Bayesian Estimation.
- 7. Hypothesis Testing and Analysis of Variance (1L):** Goodness of fit, Confidence Intervals.

Course Business

1. Professor: Floyd B. Hanson

- Email: fhanson@uchicago.edu;
- Office Hour: Mondays, 5PM (?), FinMath Lounge/Lab E7.

2. TAs: TBA

- Email: TBA
- Office Hours: TBA.
- Review Sessions: TBA.

3. Class Webpage:

FINM 345/Stat 339 Current Syllabus

or *Chalk* → Login → FINM 33100 (Winter 09) → Syllabus, etc.

4. Texts:

- These explanatory *FINM 331 Lecture Notes*;
- Optionally Recommended Text: J.R. Rice, *Math. Stat. and Data Anal.*;
- Other Texts:
 - R. Carmona, *Stat. Anal. of Fin. Data*;
 - S. Weisberg, *Appl. Lin. Regression*.

5. Recommended Computational System: **MATLAB** (class examples will be based on MATLAB):

- D. & N. Higham (brothers), *MATLAB Guide* is highly recommended, see *SIAM Books*.
- **R**, **S** and **Excel** or **Maple/Mathematica** are acceptable for assignments, but you are on your own.

6. Grading:

- **Homework (see also Syllabus):**

- There will be about 8-9 graded homework sets;
- You may consult with other student about the ideas involved;
- Submitted homework must be the individual student's own work;
- Similar solutions a problem will receive discounted grades with the problem credit divided over participants;
- Codes and/or worksheets need to be submitted with computational solutions;
- Graphical and tabular presentation needs to be professional;
- In general, problems will be graded on relative quality and not just or errors;
- Submitted assignments should be in PDF or MSword or similar file format, but using unformatted, plain text can be used within these formats.
- Penalties will be deducted for late assignment or problem submission, with negative credit for missing homework sets.

- **Exams:** There will be at least one, a take-home final exam and possible an in-class mid-term exam.
- **Final Grade:** The grade will be based upon a combination of homework and final exam, proportioned to reflect the effort involved.

Prerequisite Knowledge:

1. Introductory Probability:

- see for instance, F. B. Hanson, *Online Appendix B Preliminaries in Probability and Analysis*,
- and/or J. R. Rice, Chs. (1)-2 Probability and Random Variables, Sects. 3.1-3.5 Joint Distributions, 4.1-4.4 Expected Values.

2. Introductory Statistics, especially linear regression:

- see the review session notes, N. O. Nygaard, *Introduction to Statistics and Data Analysis*,
- and perhaps Nygaard's accompanying **Linear Algebra Notes** ;
- and/or J. R. Rice, Sects. 14.1-14.5 Linear Least Squares.

3. Very Basic MATLAB including Statistics Toolbox that comes with the Student Edition (*MATLAB Student Version*);

- See also Nygaard's review sessions previously mentioned for MATLAB techniques.

Some Related Resources of the Professor:

1. **Prior UIC course:** *Math 586 Computational Finance*, Spring 2008.
2. **Applied stochastics textbook:** F.B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation*, SIAM Books, 2007.
 - *Online Appendix C: MATLAB Programs* (listings of sample codes used to make book figures);
 - *MATLAB Source Codes Directory*, source m-files as individual files or zip-file of all m-files.
3. *Quantitative Finance References and Related References*, annotated books and links in finance and related topics.

Financial Data Analysis:

1.1 Introduction to Data and Models (with Review):

- *The model does not have much meaning without the data and the data does not have much meaning without the model;*
i.e., the model must reflect reality by being a good representation of the data, including parameter and extreme values, and the data needs a model in order to interpret the data.
- Let the financial data represent the Asset Price A (say, in dollars) and is given by a vector *sample* $\vec{A} = [A_i]_{n \times 1}$ of $n + 1$ discrete *observations* at times $\vec{t}^{(a)} = [t_i]_{(n+1) \times 1}$ in years, with spacing usually fixed with $\Delta t = t_{i+1} - t_i$, ignoring weekends and holidays.
- The basic *Asset Return* or *relative asset change* is given by

$$\vec{RR} = [R_i]_{n \times 1} = [(A_{i+1} - A_i)/A_i]_{n \times 1}, \quad (1)$$

provided $A_i \neq 0$ and is appropriate for *additive* combinations of randomness.

- However, if the randomness (uncertainty or noise) appears in *multiplicative* combinations as assumed in many financial models, then the *Log-Return* (often just called the “return”) is used,

$$\vec{LR} = [LR_i]_{n \times 1} \equiv [\log(A_{i+1}) - \log(A_i)]_{n \times 1} \quad (2)$$

$$\stackrel{\text{lol}}{=} [\log(A_{i+1}/A_i)]_{n \times 1},$$

“log”, here is the same as the natural log (i.e., “ln”) following the MATLAB usage.

- The *return and log-return are asymptotically equivalent* for small returns, since

$$LR_i = \log(A_{i+1}/A_i) = \log(1 + (A_{i+1} - A_i)/A_i) \quad (3)$$

$$\sim (A_{i+1} - A_i)/A_i = R_i$$

as $R_i \rightarrow 0$.

(Note that $f(x) \sim g(x)$ as $x \rightarrow x_0$ if $f(x)/g(x) \rightarrow 1$ as $x \rightarrow x_0$ and $f(x)$ is said to be *asymptotic* to $g(x)$; in applied math, this symbol (\sim) is **not** used for equivalent in distribution ($\stackrel{\text{dist}}{=}$).)

1.2 Independent Observations:

- Let $\vec{X} = [X_i]_{n \times 1}$ be a vector of independent (random) observations with common mean or expectation $\mu_X \equiv E_X[X_i]$ and variance $\sigma_X^2 \equiv \text{Var}_X[X_i] \equiv E_X[(X_i - \mu_X)^2]$.
{Warning: the subscript X denotes with respect to X ; common mean and variance essentially strictly excludes samples from time-dependent markets, but may be used in practice anyway.}
- Recall that X_j is *independent* of X_i if the joint distribution separates into respective univariate distributions,

$$F_{X_i, X_j}(x_i, x_j) \stackrel{\text{ind}}{=} F_{X_i}(x_i) \cdot F_{X_j}(x_j), \quad (4)$$

where $F_{X_i}(x_i) \equiv \text{Prob}[X_i \leq x_i]$ for each i and $F_{X_i, X_j}(x_i, x_j) \equiv \text{Prob}[X_i \leq x_i, X_j \leq x_j]$ in the continuous **random variable (RV)** case and similarly in the discrete RV case with distribution $F_{X_i}(k) \equiv \text{Prob}[X_i = k]$ for integer k and each i , etc.

- If the densities exist and are integrable in the continuous RV case, then the joint density also separates,

$$f_{X_i, X_j}(x_i, x_j) \stackrel{\text{ind}}{=} f_{X_i}(x_i) \cdot f_{X_j}(x_j), \quad (5)$$

where the univariate density is given by the ordinary derivative of the corresponding univariate distribution function,

$$f_{X_i}(x_i) \equiv \frac{dF_{X_i}}{dx_i}(x_i) = F'_{X_i}(x_i), \quad (6)$$

for each i and the bivariate density is given by the second partial derivative of the corresponding bivariate distribution function,

$$f_{X_i, X_j}(x_i, x_j) \equiv \frac{\partial^2 F_{X_i, X_j}}{\partial x_i \partial x_j}(x_i, x_j), \quad (7)$$

for each pair $\{i, j\}$. The separation of the densities following from the successive single derivatives: let $F(x, y) = G(x) \cdot H(y)$, then $\partial F / \partial x = G'(x) \cdot H(y)$ and $\partial^2 F / (\partial x \partial y) = G'(x) \cdot H'(y)$.

{Alert: Only one derivative is taken at a time, so partial differentiation is much like ordinary differentiation.}

- Given a density $f_X(x)$ for RV X , the **expected value, expectation or mean**, if it exists, of $g(X)$ is

$$\mathbf{E}_X[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx. \quad (8)$$

- If $\{X_1, X_2\}$ are independent RVs with density $f_{X_1}(x_1)f_{X_2}(x_2)$, then the **expectation**, if it exists, of the product $g_1(X_1)g_2(X_2)$ is

$$\begin{aligned} \mathbf{E}_{\{X_1, X_2\}}[g_1(X_1)g_2(X_2)] &= \int_{-\infty}^{\infty} g_1(x_1) f_{X_1}(x_1) dx_1 \\ &\quad \cdot \int_{-\infty}^{\infty} g_2(x_2) f_{X_2}(x_2) dx_2. \end{aligned} \quad (9)$$

- The **variance**, if it exists, of the RV X with density $f_X(x)$ is

$$\begin{aligned} \mathbf{Var}_X[X] &= \mathbf{E}_X[(X - \mathbf{E}_X[X])^2] = \mathbf{E}_X[X^2] - E_X^2[X] \\ &= \int_{-\infty}^{\infty} x^2 f_X(x) dx - E_X^2[X]. \end{aligned} \quad (10)$$

- Alternate form for the distribution of RV X is

$$F_X(x) \equiv \mathbf{Prob}[X \leq x] = \int_{-\infty}^x f_X(y) dy. \quad (11)$$

1.3 Statistics of Sample Mean of Independent Observations:

- Let X_i for $i = 1 : n$ be a sample of **independent and identically distributed (IID)** observations, with mean μ_X and variance σ_X^2 , i.e., theoretical sample observations of the some characteristic of some population.
- Define the **sample mean** of the n **discrete IID observations** be

$$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^n X_i. \quad (12)$$

Note that if the X_i are RVs, then so \bar{X}_n is a RV.

- The (simple) **sample variance** is

$$s_n^2 \equiv \overline{((X - \bar{X}_n)^2)}_n \equiv \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad (13)$$

and also is a RV.

- The **expectation of the sample mean** is, by the IID properties of X_i , as well as the linear properties of the expectation operator (\mathbf{E}),

$$\mu_{\bar{X}_n} \equiv \mathbf{E}_X [\bar{X}_n] \stackrel{\mathbf{E}}{\stackrel{\text{lin}}{=}} \frac{1}{n} \sum_{i=1}^n \mathbf{E}_X [X_i] \stackrel{\text{iid}}{=} \frac{1}{n} \sum_{i=1}^n \mu_X = \mu_X, \quad (14)$$

so \bar{X}_n is an **unbiased estimate** of the population mean μ_X , since $\mathbf{E}_X [\bar{X}_n] = \mu_X$. Let $\hat{\mu}_X = \bar{X}_n$ denote that \bar{X}_n is a proper *estimate* of μ_X .

- The **expectation of the sample variance** is,

$$\begin{aligned}
\mu_{s_n^2} &\equiv \mathbf{E}_X [s_n^2] \stackrel{\text{E}}{\stackrel{\text{lin}}{=}} \frac{1}{n} \sum_{i=1}^n \mathbf{E}_X \left[\left(X_i - \frac{1}{n} \sum_{j=1}^n X_j \right)^2 \right] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbf{E}_X \left[\left(\frac{1}{n} \sum_{j=1}^n ((X_i - \mu_X) - (X_j - \mu_X)) \right)^2 \right] \\
&\stackrel{\text{E}}{\stackrel{\text{lin}}{=}} \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \mathbf{E}_X [((X_i - \mu_X) - (X_j - \mu_X)) \\
&\quad \cdot ((X_i - \mu_X) - (X_k - \mu_X))] \tag{15} \\
&= \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \mathbf{E}_X [(X_i - \mu_X)^2 - (X_i - \mu_X)(X_j - \mu_X) \\
&\quad - (X_i - \mu_X)(X_k - \mu_X) + (X_j - \mu_X)(X_k - \mu_X)] \\
&\stackrel{\text{iid}}{=} \frac{(n^3 - n^2 - n^2 + n^2)}{n^3} \sigma_X^2 = \frac{n-1}{n} \sigma_X^2,
\end{aligned}$$

so s_n^2 is a ***biased estimate*** of the population variance σ_X^2 since $\mathbf{E}_X [s_n^2] \neq \sigma_X^2$, but the bias is small for n large.

However, it is easy to see that

$$\hat{s}_n^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad (16)$$

is an **unbiased estimate** of the population variance σ_X^2 since $\mathbf{E}_X [\hat{s}_n^2] = \sigma_X^2$. Division by $(n-1)$ instead of n can be seen as the loss of one **degree of freedom (DOF)** because $\mathbf{E}[X_i - \bar{X}_n] = 0$.

- The *variance of the sample mean* is similarly

$$\begin{aligned} \sigma_{\bar{X}_n}^2 &\equiv \text{Var}_X [\bar{X}_n] \equiv \mathbf{E}_X \left[(\bar{X}_n - \mu_{\bar{X}_n})^2 \right] \\ &= \mathbf{E}_X \left[\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu_X) \right)^2 \right] \\ &\stackrel{\mathbf{E}}{\stackrel{\text{lin}}{=}} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathbf{E}[(X_i - \mu_X)(X_j - \mu_X)] \\ &\stackrel{\text{iid}}{=} \frac{1}{n^2} \sum_{i=1}^n \sigma_X^2 = \frac{\sigma_X^2}{n} \rightarrow 0^+ \text{ as } n \rightarrow 0^+, \end{aligned} \quad (17)$$

i.e., as n becomes very large the variance of $\sigma_{\bar{X}_n}^2$ becomes very small.

The reduction in the sample mean variance in Eq. (17) is the basis for the probable error in the powerful **Monte Carlo random sample simulation approximations**, since this error is the order of the standard deviation

$$\sigma_{\bar{X}_n} = \frac{\sigma_X}{\sqrt{n}} \quad (18)$$

and is called the **standard error (SE) of \bar{X}_n** , yielding a significant reduction from the population variance σ_X^2 to the sample mean variance $\sigma_{\bar{X}_n}^2$ by a factor of $1/n$. However, n has to be very large, e.g., changing from $n = 100$ to $n' = 1,000,000$, then $\sqrt{n'/n} = 100$ which means only a **1% reduction** in the standard error of \bar{X}_n for **10,000** times more computational effort.

1.4 Statistics of Discrete Additive Financial Models:

- A simple discretized version of Louis **Bachelier**'s (1900 thesis under Henri Poincaré) asset price model in the option pricing problem is the **affine recursion with purely additive uncertainties** for new asset price A_{i+1} ,

$$A_{i+1} = A_i + \Delta A_i = A_i + \mu \Delta t + \sigma \sqrt{\Delta t} \cdot Z_i, \text{ given } A_1, \quad (19)$$

for $i = 1:n$ fixed time-steps $\Delta t = t_{i+1} - t_i$, where Z_i is IID with mean $\mathbf{E}[Z_i] = \mu_Z$ and variance $\text{Var}[Z_i] = \sigma_Z^2$ random variable or uncertainty, but since there are already similar coefficients $\{\mu, \sigma\}$, let Z_i be zero-mean $\mu_Z = 0$ and unit-variance $\sigma_Z^2 = 1$) (It is often assumed that Z_i is IID normal.) Hence, $\mathbf{E}[\Delta A_i] = \mu \Delta t$ is the mean increase in A_i in time-step Δt and $\mathbf{E}[\Delta A_i] = \sigma^2 \Delta t$ is its variance.

- In **real markets over long periods**, $\{\mu, \sigma, \Delta t\}$ are not likely to be time-independent. Let $\{\mu, \sigma, \Delta t\}$ be independent of the price A_j or noise Z_j at any time t_j . Also, let $|Z_i| \leq B_z$, i.e., “*practically bounded*” noise, and $|\mu \Delta t + \sigma \sqrt{\Delta t} Z_i| \leq B_{\Delta A}$ to ensure bounded prices. Typically, $\Delta t \simeq 1/252$ is one trading day in year units.

- Summing the above additive recursion yields for $i = 1:n$,

$$A_{i+1} = A_1 + \sum_{j=1}^i (\mu\Delta t + \sigma\sqrt{\Delta t} \cdot Z_j), \quad (20)$$

noting again that if the Z_i are **RVs**, then so is A_{i+1} for $i = 1:n$.

- The new *asset mean* is

$$\begin{aligned} \mu_{A_{i+1}} &\equiv \mathbf{E}_Z [A_{i+1}] \stackrel{\text{E}}{\text{lin}} A_1 + \sum_{j=1}^i (\mu\Delta t + \sigma\sqrt{\Delta t} \mathbf{E}_Z [Z_j]) \\ &= A_1 + i(\mu\Delta t + \sigma\sqrt{\Delta t} \mu_Z) = A_1 + i\mu\Delta t, \end{aligned} \quad (21)$$

given $\mu_Z = 0$, *growing linearly* with time-step count i , as a Malthusian (1768) growth of an agricultural commodity, rather than a population or a financial asset.

{Alert: Since $\mu_{A_i} = A_1 + (i-1)\mu\Delta t$ is dependent on the time-step index i , then the A_i are RVs but not IID RVs, although ΔA_i are IID RVs.}

- The *asset variance* is

$$\begin{aligned}
\sigma_{A_{i+1}}^2 &\equiv \text{Var}_Z [A_{i+1}] \equiv \mathbf{E}_Z \left[(A_{i+1} - \mu_{A_{i+1}})^2 \right], \\
&= \mathbf{E}_Z \left[\left(\sum_{j=1}^i \sigma \sqrt{\Delta t} (Z_j - \mu_Z) \right)^2 \right] \\
&\stackrel{\text{E}}{\stackrel{\text{lin}}{=}} \sigma^2 \Delta t \sum_{j=1}^i \sum_{k=1}^i \mathbf{E}_Z [Z_j Z_k] \\
&\stackrel{\text{iid}}{=} \sigma^2 \Delta t \sum_{j=1}^i \sum_{k=1}^i \delta_{j,k} \sigma_Z^2 = \sigma^2 \Delta t \sum_{j=1}^i \sigma_Z^2 \\
&= i \sigma^2 \Delta t,
\end{aligned} \tag{22}$$

given $\sigma_Z^2 = 1$ and where $\delta_{i,j}$ is the discrete, **Kronecker delta**, which is **1** when $i = j$, else **0**.

{Alert: $(\sum_i^n a_i)^2 = \sum_i^n \sum_j^n a_i a_j \neq \sum_i^n a_i^2$,
e.g., $(a_1 + a_2)^2 = a_1^2 + 2a_1 a_2 + a_2^2$.}

- Summarizing in financial return terms, the i th **mean return** starting at time index i for $i = 1 : n$ is

$$\mu_{R_i} \equiv \mathbf{E}_Z[R_i] \equiv \mathbf{E}_Z[\Delta A_i] = \mu \Delta t. \quad (23)$$

- Also, the i th **return variance** starting at time index i is

$$\sigma_{R_i}^2 \equiv \text{Var}_Z[R_i] = \mathbf{E}_Z[(R_i - \mu_{R_i})^2] \stackrel{\text{iid}}{=} \sigma^2 \Delta t. \quad (24)$$

- *The relative change form of the return is appropriate for the additive model, where uncertainties are additive.*

1.5 Statistics of Discrete Multiplicative Financial Models:

- A simple, **discretized version of Black-Scholes and Merton's asset price model in the option pricing problem** (see Spring 1973 companion model and mathematical justification papers) is the **linear recursion with purely multiplicative factor uncertainties** for the option prices,

$$A_{i+1} = A_i + \Delta A_i = A_i(1 + \mu\Delta t + \sigma\sqrt{\Delta t} \cdot Z_i), \quad (25)$$

for $i = 1 : n$, given A_1 , where the terms otherwise satisfy the same conditions as in the previous additive noise section.

- Solving the above recursion in terms of products rather than sums (*note that taking logarithms turns the solution into the additive form*) yields for $i = 1 : n$,

$$A_{i+1} = A_1 \prod_{j=1}^i (1 + \mu\Delta t + \sigma\sqrt{\Delta t} \cdot Z_j). \quad (26)$$

{Alert: If the Z_i are RVs, then so is ΔA_i , but A_i is only a RV and not IID, since the basic statistics vary with i .}

- The new *asset mean* in this multiplicative case is

$$\begin{aligned}\mu_{A_{i+1}} &\equiv \mathbf{E}_Z [A_{i+1}] \stackrel{\text{iid}}{=} A_1 \prod_{j=1}^i (1 + \mu\Delta t + \sigma\sqrt{\Delta t}\mu_Z) \\ &= A_1 (1 + \mu\Delta t)^i,\end{aligned}\tag{27}$$

using the zero-mean property of the Z_i , where $\prod_{j=1}^i$ is the product symbol from **1** to i . Note that the asset grows like *discretely compounded interest* which it should, considering *geometric Brownian motion* of the asset.

- The *asset variance* for multiplicative noise is similarly

$$\begin{aligned}\sigma_{A_{i+1}}^2 &\equiv \text{Var}_Z [A_{i+1}] \equiv \mathbf{E}_Z \left[(A_{i+1} - \mu_{A_{i+1}})^2 \right] \\ &= A_1^2 \mathbf{E}_Z \left[\left(\prod_{j=1}^i \sigma\sqrt{\Delta t} Z_j \right)^2 \right] \\ &= A_1^2 \prod_{j=1}^i \sigma\sqrt{\Delta t} \prod_{k=1}^i \sigma\sqrt{\Delta t} \mathbf{E}_Z [Z_j Z_k] \\ &\stackrel{\text{iid}}{=} A_1^2 (\sigma^2 \Delta t)^i \prod_{j=1}^i \prod_{k=1}^i \delta_{j,k} = A_1^2 (\sigma^2 \Delta t)^i,\end{aligned}\tag{28}$$

where $\delta_{j,k}$ is again the **Kronecker delta**.

{Alert: $(\prod_{j=1}^i a_j)^2 = \prod_{j=1}^i a_j \prod_{k=1}^i a_k$ and $\prod_{j=1}^i c = c^i$.}

- The i th **mean log-return** starting at time index i for $i = 1 : n$ is

$$\begin{aligned} \mu_{LR_i} &\equiv E_Z [\log(A_{i+1}/A_i)] = E_Z [\log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i)] \\ &\stackrel{\text{iid}}{=} \mu_{\log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z)}. \end{aligned} \quad (29)$$

- The i th **log-return variance** starting at time index i is

$$\begin{aligned} \sigma_{LR_i}^2 &\equiv \text{Var}_Z [LR_i] = E_Z [(LR_i - \mu_{LR_i})^2] \\ &= E_Z \left[(\log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i) - \mu_{\log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z)})^2 \right] \\ &\stackrel{\text{iid}}{=} \sigma_{\log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z)}^2. \end{aligned} \quad (30)$$

The identically distributed (ID of IID) is required for the logarithm.

- ***The relative change form of the return is appropriate for the multiplicative model, where uncertainties are multiplicative.***

- The **multiplicative model log-form** has an additive form, for $i = 1 : n$

$$\log(A_{i+1}) = \log(A_i) + \log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i). \quad (31)$$

- Solving yields,

$$\log(A_{n+1}) = \log(A_1) + \sum_{i=1}^n \log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i). \quad (32)$$

- When the time-stop Δt is **small** ($\Delta t \ll 1$), the **log-asset** has the asymptotic (\sim symbol) behavior, with error $\mathbf{o}(\Delta t)$, i.e., smaller than order Δt ,

$$\log(A_{n+1}) \sim \log(A_1) + \sum_{i=1}^n (\sigma\sqrt{\Delta t}Z_i + (\mu - \sigma^2 Z_i^2 / 2)\Delta t), \quad (33)$$

since $\log(1 + \varepsilon) = \varepsilon - \varepsilon^2 / 2 + \mathbf{O}(\varepsilon^3) \sim \varepsilon - \varepsilon^2 / 2$ for $\varepsilon = \sqrt{\Delta t} \ll 1$.

*{Alert: Asymptotic **big and little Oh**: $f(x) = \mathbf{O}(g(x))$ if $f(x)/g(x)$ is bounded as $x \rightarrow x_0$; $f(x) = \mathbf{o}(g(x))$ or $f(x) \ll g(x)$ if $f(x)/g(x) \rightarrow 0$ as $x \rightarrow x_0$.}*

- Similarly, the i th **log-return** behaves as, with error $o(\Delta t)$,

$$\text{LR}_i \sim \log(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i) = \sigma\sqrt{\Delta t}Z_i + (\mu - \sigma^2 Z_i^2/2)\Delta t. \quad (34)$$

- The i th **mean log-return** behaves as,

$$\mathbf{E}_Z[\text{LR}_i] \sim (\mu - \sigma^2/2)\Delta t, \quad (35)$$

since $\text{Var}_X[X] = \mathbf{E}_X[X^2] - (\mathbf{E}_X[X])^2$ or $\sigma_X^2 = \mu_{X^2} - \mu_X^2$.

{Alert for 345 veterans: Result (35) confirmed the corresponding stochastic calculus result for continuous models without the stochastic calculus.}

- The i th **log-return variance** behaves as, again neglecting terms smaller than order Δt ,

$$\begin{aligned} \text{Var}_Z[\text{LR}_i] &\sim \mathbf{E}_Z[(\sigma\sqrt{\Delta t}Z_i - \sigma^2(Z_i - 1)\Delta t/2)^2] \\ &\sim \mathbf{E}_Z[\sigma^2\Delta tZ_i^2] = \sigma^2\Delta t, \end{aligned} \quad (36)$$

so for sufficiently small time-steps $\Delta t \ll 1$, the **log-return** LR_i of the multiplicative model will be **asymptotically approximate IID**, neglecting terms smaller than order Δt , i.e., with error $o(\Delta t)$.

1.6 Stochastic (Random) Nature of Financial Data:

Fig01: Adj. Closing Histogram, '88-'08

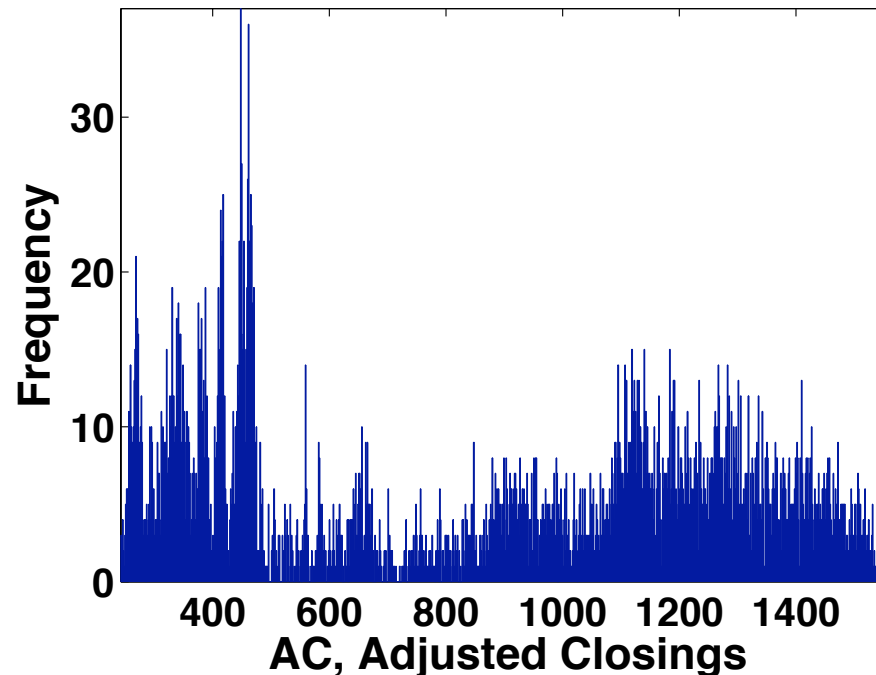


Figure 3: **S&500** Index **Daily Adjusted Closings** $AC(t)$ for $t=1:NAC$, from 1988 to 2008 (post-1987) showing scattered behavior of the price **without any recognizable probability distribution seen, certainly not a normal distribution.** (aka Fig01; $NAC=AC\text{-count.}$)

Fig02: Abs. Return Histogram, '88-'08

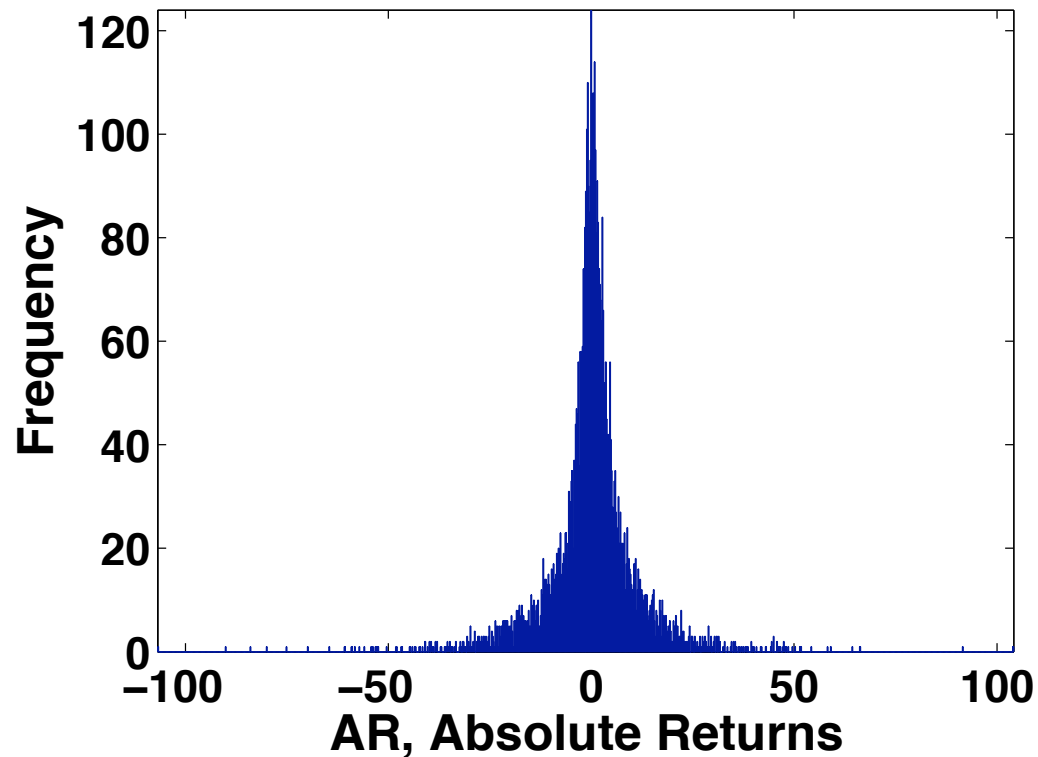


Figure 4: **S&500** Index Daily **Absolute Returns or Differences** $AR(t) = \widetilde{AR}(t+1) = AC(t+1) - AC(t) \equiv \Delta AC(t)$ for $t=1:NAC-1$, from 1988 to 2008 (post-1987) showing **more central behavior**, resembling a **very narrow normal distribution (or spike)** with many discrete deviations from the normal. (**aka Fig02.**)

Fig03: Rel. Return Histogram, '88-'08

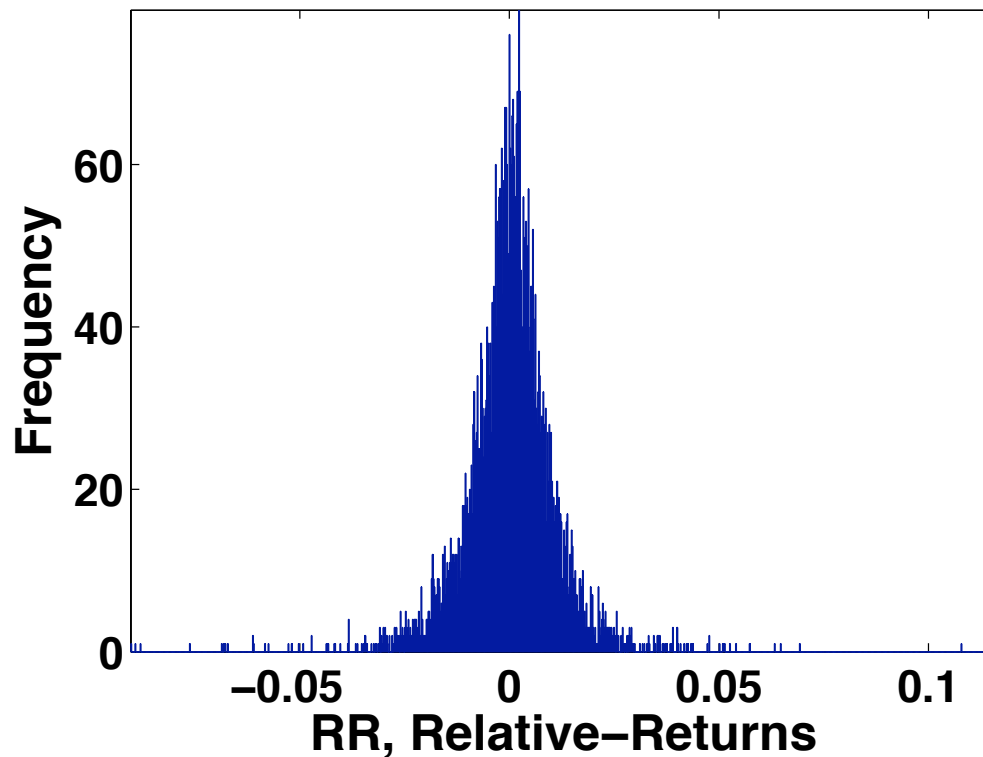


Figure 5: **S&500** Index Daily **Relative Returns** $RR(t) = AC(t + 1)/AC(t) - 1 \equiv \Delta AC(t)/AC(t)$ for $t=1:NAC-1$, from 1988 to 2008 (post-1987) showing a **more developed normal distribution with wider spread due to reduction of the scale** of the returns **and many rare tail events.** (aka **Fig03.**)

Fig04: LR Histogram Vs. NormPDF+1, '88-'08

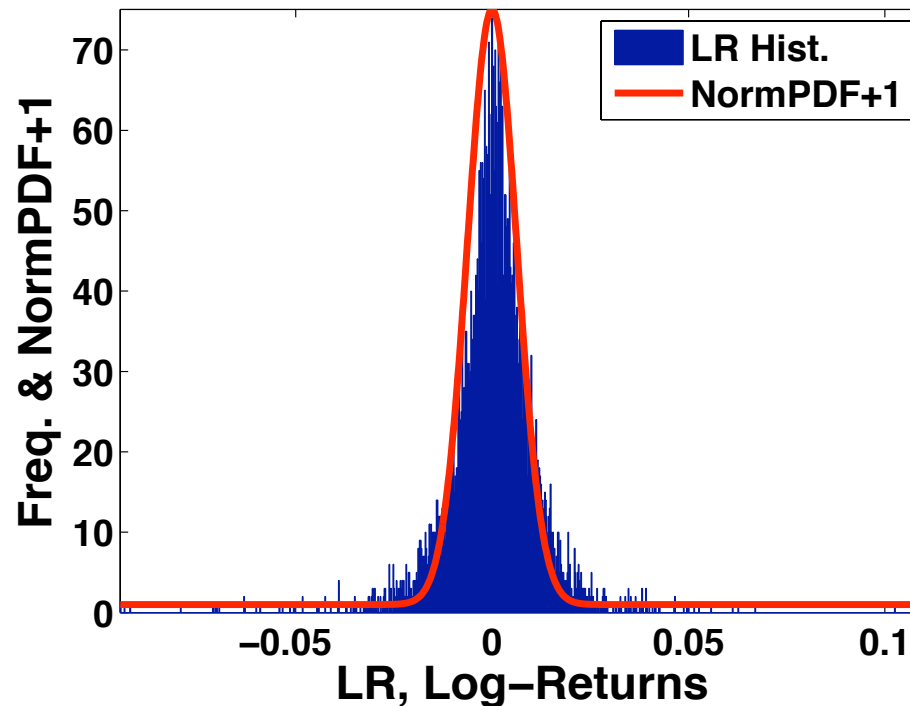


Figure 6: **S&500** Index Daily **Log Returns** $LR(t) = \log(AC(t+1)) - \log(AC(t)) = \log(1+RR(t)) \sim RR(t)$ for $t=1:NAC-1$ & $RR(t) \ll 1$, from 1988 to 2008 (post-1987) showing wide spread and tail event behavior similar to $RR(t)$. In **red**, an approximate normal density is overlaid plus a unit to account for fat tails from jump of crashes and bubbles. **So, the infinite normal tails have little probability compared to jumps.**

1.7 Law of Large Numbers (LLN; Weak Form):

- **Theorem 1.6. Weak Law of Large Numbers:**

Let $\vec{X} = [X_i]_{n \times 1}$ be a sample vector of **independent RVs with identical mean $\mu_X = E[X_i]$ and variance $\sigma_X^2 = \text{Var}[X_i]$** . Let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ be the **sample mean**, then for any fixed $R > 0$,

$$\text{Prob}[|\bar{X}_n - \mu_X| > R] \rightarrow 0^+ \text{ as } n \rightarrow \infty, \quad (37)$$

i.e., converges to **0** in probability.

- **Proof:** A simple proof depends on the variance and follows from **Chebyshev's inequality**, although stronger proofs do not, i.e.,

$$\text{Prob}[|X - \mu_X| > R] \leq \sigma_X^2 / R^2, \quad (38)$$

so since $E[\bar{X}_n] = \mu_X$ and $\text{Var}[\bar{X}_n] = \sigma_X^2 / n$, then

$$\text{Prob}[|\bar{X}_n - \mu_X| > R] \leq \frac{\sigma_X^2}{nR^2} \rightarrow 0^+ \text{ as } n \rightarrow \infty. \quad (39)$$

- The proof of Chebyshev's inequality is as an exercise. However, a more useful form in finance is, for any $K > 0$,

$$\text{Prob}[|\bar{X}_n - \mu_X| > K\sigma_X] \leq \frac{1}{nK^2} \rightarrow 0^+ \text{ as } n \rightarrow \infty, \quad (40)$$

i.e., the risk of a large change (crash or rally) of K standard deviations or greater in the sample mean has probability

$O(1/(nK^2))$.

- Markov's simpler inequality can also be used for Chebyshev's inequality and it states:

Theorem 1.6M Markov's inequality: Let X be a RV such that $\text{Prob}[X \geq 0] = 1$ and $\mathbf{E}[X]$ exists, then

$$\text{Prob}[X \geq x] \leq \mathbf{E}_X[X]/x = \mu_X/x. \quad (41)$$

- ***Proof Sketch of Markov's Inequality:*** A quick formal justification for the case that X has a well-defined density $f_X(x)$ on $[0, \infty)$ line implies,

$$\begin{aligned} E_X[X] &= \int_0^{\infty} y f_X(y) dy = \left(\int_0^{y < x} + \int_{y \geq x}^{\infty} \right) y f_X(y) dy \\ &\geq \int_{y \geq x}^{\infty} y f_X(y) dy \geq x \int_{y \geq x}^{\infty} f_X(y) dy = x \text{Prob}[X \geq x], \end{aligned} \tag{42}$$

rearranging yields the Markov inequality.

1.8 Central Limit Theorem (CLT) and Asymptotic Normal Distributions:

- **Theorem 1.7. Central Limit Theorem :**

Let $\vec{X} = [X_i]_{n \times 1}$ be a sample vector of *independent RVs* and *there exists an identical mean* $\mu_X = \mathbb{E}[X_i]$ and *variance*

$\sigma_X^2 = \text{Var}[X_i]$, assuming the first three moments are bounded. Let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ be the *sample mean*, then for any fixed x in $(-\infty, +\infty)$,

$$\lim_{n \rightarrow \infty} \text{Prob} \left[\frac{\bar{X}_n - \mu_X}{\sigma_X / \sqrt{n}} \leq x \right] = \Phi(x), \quad (43)$$

where Φ is the standard (mean zero, unit variance) normal distribution function.

- The central limit theorem or CLT says that the scaled sample mean converges in probability to the standard normal distribution,

$$\Phi(x) = F_X^{(n)}(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy. \quad (44)$$

- **Proof of CLT:** The proof is by **moment generating functions** (see Rice (2007, p. 184; *but a mini-tutorial is included in proof*). There are many versions of the CLT and some of the strong CLTs relax some of the necessary conditions.) For the IID RV with density $f_X(x)$, the moment generating function is an exponential transformation using a transform variable, say t , so for the observational RV,

$$\begin{aligned}
 M_{X_i}(t) &\equiv \mathbf{E}_{X_i} [e^{tX_i}] \stackrel{\text{iid}}{=} M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\
 &= \mathbf{E}_{X_i} [e^{tX_i}] = e^{t\mu_X} \mathbf{E}_X [e^{t(X-\mu_X)}] \\
 &= e^{t\mu_X} \mathbf{E}_X [1 + 0.5t^2\sigma_X^2 + O(t^3)],
 \end{aligned} \tag{45}$$

for all $i = 1:n$ and upon the convenient Taylor expansion of the moment generation function for the deviation $\delta X_i \equiv X_i - \mu_X$ when t is sufficiently small. Note the identity $M_{\delta X_i}(t) = e^{-t\mu_X} M_{X_i}(t)$.

Similarly, since also $\mu_{\bar{X}_n} = \mu_X$, the sample mean deviation is $\delta\bar{X}_n \equiv \bar{X}_n - \mu_X = \overline{(X - \mu_X)}_n$. Then,

$$\begin{aligned}
 M_{\delta\bar{X}_n}(t) &\equiv \mathbf{E}_X \left[e^{t\delta\bar{X}_n} \right] = \mathbf{E}_X \left[e^{\frac{t}{n} \sum_{i=1}^n \delta X_i} \right] \\
 &\stackrel{\text{loe}}{=} \mathbf{E}_X \left[\prod_{i=1}^n e^{\frac{t}{n} \delta X_i} \right] \stackrel{\text{iid}}{=} \prod_{i=1}^n \mathbf{E}_X \left[e^{\frac{t}{n} \delta X} \right] \\
 &\stackrel{\text{prod}}{\underset{\text{taylor}}{=}} \mathbf{E}_X^n \left[1 + \frac{t}{n} \delta X + 0.5 \left(\frac{t}{n} \delta X \right)^2 + \mathbf{O}^3 \left(\frac{t}{n} \right) \right] \\
 &= \left(1 + 0.5 \left(\frac{t}{n} \sigma_X \right)^2 + \mathbf{O}^3 \left(\frac{t}{n} \right) \right)^n \tag{46} \\
 &\stackrel{\text{bin}}{=} \left(1 + 0.5 \frac{(t\sigma_X)^2}{n} + \mathbf{O} \left(\frac{t^3}{n^2} \right) \right) \\
 &\sim \left(1 + 0.5 \left(\frac{t\sigma_X}{\sqrt{n}} \right)^2 \right) \sim e^{0.5(t\sigma_X)^2/n} \stackrel{\text{cts}}{=} M_{\delta X^{(n)}} \left(\frac{t}{\sqrt{n}} \right),
 \end{aligned}$$

as $n \rightarrow \infty$ for fixed t , where $\delta X^{(n)} = X^{(n)} - \mu_X$ is the deviate normal RV with distribution $F_X^{(n)}(x; 0, \sigma_X^2/n)$. Hence,

$\delta\bar{X}_n \xrightarrow{\text{dist}} F_X^{(n)}(x; 0, \sigma_X^2/n)$ as $n \rightarrow \infty$.

QED

- In the CLT proof, the **complete the square (CTS)** technique was used to evaluate the moment generating function or the expectation of an exponential of a normal variate, in particular by completing the square of the exponent as follows:

$$\begin{aligned}
 M_{\delta X^{(n)}}(t) &\equiv \mathbf{E}_{\delta X^{(n)}} \left[e^{t\delta X^{(n)}} \right] \stackrel{\text{normal}}{=} \int_{-\infty}^{\infty} e^{tx} e^{-0.5x^2/\sigma_X^2} \frac{dx}{\sqrt{2\pi\sigma_X^2}} \\
 &\stackrel{\text{lcd}}{=} \int_{-\infty}^{\infty} e^{-0.5(x^2 - 2\sigma_X^2 tx)/\sigma_X^2} \frac{dx}{\sqrt{2\pi\sigma_X^2}} \\
 &\stackrel{\text{cts}}{=} e^{0.5(t\sigma_X)^2} \int_{-\infty}^{\infty} e^{-0.5(x - t\sigma_X^2)^2/\sigma_X^2} \frac{dx}{\sqrt{2\pi\sigma_X^2}} \\
 &= e^{0.5(t\sigma_X)^2} \int_{-\infty}^{\infty} e^{-0.5y^2/\sigma_X^2} \frac{dy}{\sqrt{2\pi\sigma_X^2}} = e^{0.5(t\sigma_X)^2}.
 \end{aligned} \tag{47}$$

Hence, $M_{X^{(n)}}(t) = \exp(t\mu_X + 0.5(t\sigma_X)^2)$. In this item, technically, we should use $\{\mu_{X^{(n)}}, \sigma_{X^{(n)}}^2\}$ for basic moments.

- Also used was an **asymptotic version of the binomial theorem**,

$$\begin{aligned}
 \left(1 + \frac{a_2}{n^2} + \frac{a_3}{n^3} + \dots\right)^n &\stackrel{\text{bin}}{=} \left(1 + n \left(\frac{a_2}{n^2} + \frac{a_3}{n^3}\right) + \frac{n(n-1)}{2!} \left(\frac{a_2}{n^2}\right)^2 + \dots\right) \\
 &\stackrel{\text{O}}{=} \left(1 + \frac{a_2}{n} + \text{O}\left(\frac{1}{n^2}\right)\right).
 \end{aligned} \tag{48}$$

Note: the $a_1/n \equiv 0$ for the deviation form, else too complicated.

- **Critical Discussion of CLT (43):**

- The CLT is an asymptotic theorem that *depends on the scaling of the IID RVs* and the normal distribution results may not be seen without the right scaling. The scaling of the sample mean in (43) is the naturally scaled sample mean

$$Z_n = \frac{\bar{X}_n - \mu_{\bar{X}_n}}{\sigma_{\bar{X}_n}}, \quad (49)$$

where $\mu_{\bar{X}_n}$ is given in (14) and $\sigma_{\bar{X}_n}^2$ is given in (17).

- The *normal distribution is a good representation of the central bulk of large samples of financial data*, but a sociological finance point of view.

- In the case that the **identically distributed (ID) part of the conditions for the CLT are relaxed**, i.e., each observation RV X_i has its own basic statistics, $\mu_{X_i} = \mathbf{E}[X_i]$ and $\sigma_{X_i}^2 = \text{Var}[X_i]$, so that $\mu_{\bar{X}_n} = \sum_{i=1}^n \mu_{X_i}$ and $\sigma_{\bar{X}_n}^2 = \sum_{i=1}^n \sigma_{X_i}^2$, making the CLT still valid in for the scaled variable in (49).
- The CLT (43) normal distribution limit is requires an infinite domain $X \in (-\infty, +\infty)$, but this is an approximation, since clearly the ***world's financial wealth is bounded***, perhaps by hundreds of trillion or quadrillion dollars, and finite so infiniteness is realistically doubtful unless it is a good approximation.

- The fact that both μ_X and σ_X moments are unknown limits the usefulness of CLT, but a *practical convergence in distribution* (\xrightarrow{d}) *version of the CLT* is

$$\frac{\bar{X}_n - \mu_X}{\hat{s}_n^2[X]/\sqrt{n}} \xrightarrow{d} F_X^{(n)}(x; 0, 1), \quad (50)$$

where

$$\hat{s}_n^2[X] = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \quad (51)$$

is the unbiased estimate of σ_X^2 since $\hat{s}_n^2[X] \rightarrow \sigma_X^2$ as $n \rightarrow \infty$. This CLT form is better for determining *confidence intervals* for estimating μ_X .

- For example, suppose that the CLT is a reasonable approximation to the distribution sample mean \bar{X}_n an IID sample of n RVs with mean μ_X and variance σ_X^2 . We wish to find the probability that the sample mean is within a distance c from the population mean μ_X . We have

$$\begin{aligned} \text{Prob}[|\bar{X}_n - \mu_X| \leq c] &= \text{Prob}\left[\frac{\sqrt{n}|\bar{X}_n - \mu_X|}{\sigma_X} \leq \frac{\sqrt{nc}}{\sigma_X}\right] \\ &\stackrel{\text{clt}}{\simeq} F_X^{(n)}\left(+\frac{\sqrt{nc}}{\sigma_X}; 0, 1\right) - F_X^{(n)}\left(-\frac{\sqrt{nc}}{\sigma_X}; 0, 1\right) \quad (52) \\ &\stackrel{\text{sym}}{=} 1 - 2F_X^{(n)}\left(-\frac{\sqrt{nc}}{\sigma_X}; 0, 1\right) = p_n(c), \end{aligned}$$

where $p_n(c)$ is probability of the **confidence interval** corresponding to the distance c from the mean for a sample size of n . There are several problems options given just σ_X^2 . If $\{c, n\}$ are given, then $p_n(c)$ can be estimated using the cumulative normal **p=normcdf(x, mu, sigma)** or if $\{c, p_n(c)\}$ are given, then the sample size can be estimated using the inverse **x=norminv(p, mu, sigma)** in **MATLAB**.

- **Undervalued Extreme Rare Events Hidden in the Sample Mean:** The central part of returns is relatively dense compared to the rare events of the highest values and so are underweighted in the simple counting of the sample mean, $\bar{X}_n = \sum_{i=1}^n X_i/n$. Let the samples be decomposed into central $X_i^{(c)}$ values, isolated, rare positive values $X_j^{(r+)}$ and isolated, rare negative $X_\ell^{(r-)}$ values, i.e., X_i decomposed such that

$$\bar{X}_n = \frac{1}{n} \left(\sum_{i=1}^{n-k^{(r+)}-k^{(r-)}} X_i^{(c)} + \sum_{j=1}^{k^{(r+)}} X_j^{(r+)} + \sum_{\ell=1}^{k^{(r-)}} X_\ell^{(r-)} \right).$$

For instance, an estimate of the central part is the integral,

$$\mu_X^{(c)} = \int_{-\infty}^{\infty} x f_X^{(c)}(x) dx,$$

which is small, $2.4e-4$ for the log-returns of the 1988-2008 S&P500 Index, but an estimate of the loss part is extremely small,

$$\bar{k}^{(r-)} \max_{\ell} [|X_{\ell}^{(r-)}|] / n \simeq 1.9e-7,$$

where from the same data $\bar{k}^{(r-)} \simeq 0.0104$ jump count, $\max_{\ell} [|X_{\ell}^{(r-)}|] \simeq 0.095$ extreme log-return and $n = 5296$ log-return values. Note, in this example the observation, corresponding to the application of the CLT, is the log-return $X_i = LR_i = \log(A_{i+1}/A_i)$ using the multiplicative model (31) and $\bar{X}_n = \log(A_{n+1}/A_1)/n$ the asymptotic total average log-price change using (33).

- A *critical disadvantage of the normal distribution* representation is that it has *exponentially thin tails* and the density is order of

$$e^{-(x - \mu_X)^2 / (2\sigma_X^2)} \ll e^{-|x|} \ll \frac{1}{|x|^m} \ll 1 \quad (53)$$

as $|x| \rightarrow \infty$, but this is in contrast to the data which show that the tails are much fatter, i.e., *fat tails are observed in real markets*. The main problem is that the ***DOWNSIDE RISK*** of the market is estimated by probability of the larger losses (including crashes), i.e., for estimating $\text{Prob}[-X > R]$, which would be highly unrealistic using a normal distribution. The tail part of the distribution is usually *nonnormal*. The estimation of gains from bubbles or rallies are not as critical since they are bonanzas, except for weak accuracy of the profit forecast.

1.9 Monte Carlo Method Application:

- In many financial applications integral approximations are needed, such as expectation integrals applied for real, continuous, IID RVs, $\vec{X} = [X_i]_{n \times 1}$, with proper distribution $F_X(x) = \text{Prob}[X_i \leq x]$ and proper probability density $f_X(x)$ on $[a, b]$ (e.g., if fully infinite, then $(-\infty, +\infty)$):

$$\begin{aligned}\mu_{g(X)} &\equiv \mathbf{E}_X[g(X)] \equiv \int_a^b g(y) f_X(y) dy \\ &\simeq \frac{1}{n} \sum_{i=1}^n g(X_i) \equiv \overline{g(X)}_n,\end{aligned}\tag{54}$$

i.e., approximated by the sample mean, which is a RV with the X_i and all lower moments of the $\overline{g(X)}_n$ are assumed to exist.

- The **expected value of $\overline{g(X)}_n$** is an unbiased estimate of $\mu_{g(X)}$, i.e.,

$$\mu_{\overline{g(X)}_n} \equiv \mathbf{E}_X[\overline{g(X)}_n] \stackrel{\text{iid}}{=} \frac{1}{n} \sum_{i=1}^n \mu_{g(X)} = \mu_{g(X)},\tag{55}$$

noting that $\frac{1}{n} \sum_{i=1}^n$ is not an RV and \mathbf{E}_X is a linear operator.

- The **variance of $\overline{g(X)}_n$** is given by the variance of the sample mean formula (17) in Sect. 1.3,

$$\sigma_{\overline{g(X)}_n}^2 = \sigma_{g(X)}^2/n, \quad (56)$$

so the standard error or standard deviation is

$$\text{SE}[\overline{g(X)}_n] = \sigma_{\overline{g(X)}_n} = \sigma_{g(X)}/\sqrt{n}.$$

- However, $\sigma_{g(X)}$ is the target of the method and is *a priori* unknown, but for a probabilistic estimate of the error the **unbiased sample variance $g(X)$** can be used, i.e.,

$$\hat{s}_n^2 = \frac{1}{n-1} \sum_{i=1}^n \left(g(X_i) - \overline{g(X)}_n \right)^2. \quad (57)$$

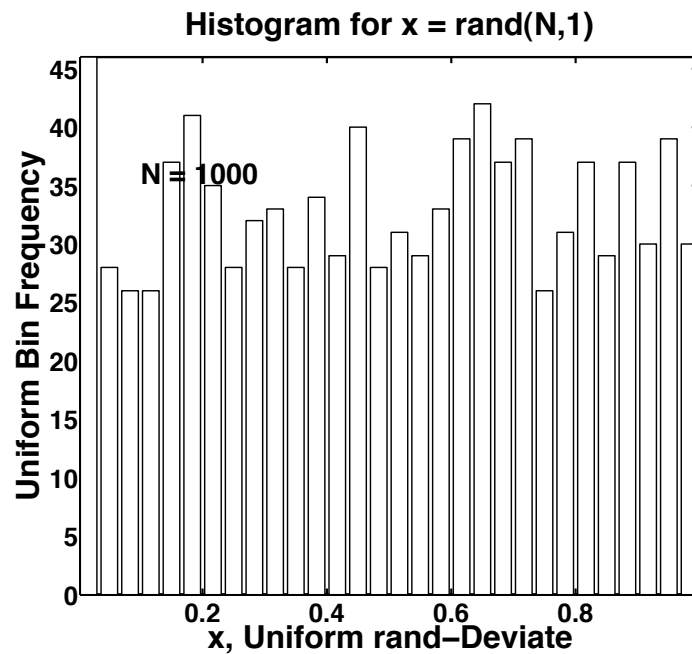
Hence, our estimate of the order of the Monte Carlo probabilistic error is the estimated **SE**,

$$\widehat{\text{SE}}_n[\overline{g(X)}] = \frac{\hat{s}_n}{\sqrt{n}}. \quad (58)$$

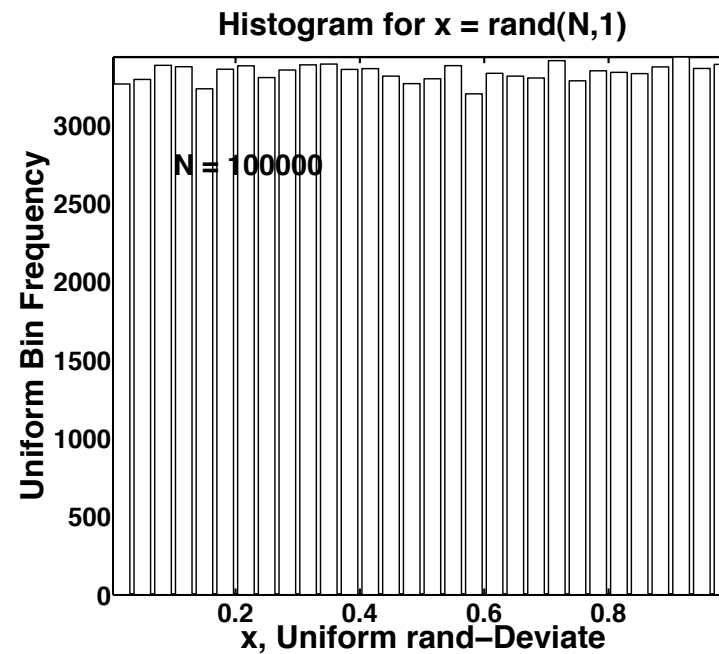
- Often, the Monte Carlo estimate, itself, is written very loosely as “ $\mu_{g(X)} = \overline{g(X)}_n \pm \hat{s}_n / \sqrt{n}$ ”, but the **order of one standard deviation margin of error** should be understood, i.e.,

$$\mu_{g(X)} = \overline{g(X)}_n + O(\hat{s}_n / \sqrt{n}).$$
- Hence, an Monte Carlo estimate can be computed by using an appropriate pseudorandom number generator like **rand(1, n)** for the row **n**-vector, **standard uniform distribution** on $[\delta, 1 - \delta] \simeq (0^+, 1^-)$ in basic **MATLAB** or **unifrnd** in the **Statistics Toolbox**, where $\delta = \text{eps}/2$ and **eps** is the machine epsilon in MATLAB, and **randn(1, n)** for the **mean zero, unit variance, standard normal distribution** on $(-\infty, +\infty)$ in basic **MATLAB** or **normrnd(0, 1, 1, n)** in the **Statistics Toolbox**, for instance. In general, **unifrnd(a, b, m, n)** \Leftrightarrow **a + (b-a)*rand(m, n)** and **normrnd(mu, sigma, m, n)** \Leftrightarrow **mu + sigma*randn(m, n)** for **m X n** samples.

{Mileage will vary with other computational math/stat. systems.}



(a) Sample size $N = 10^3$.

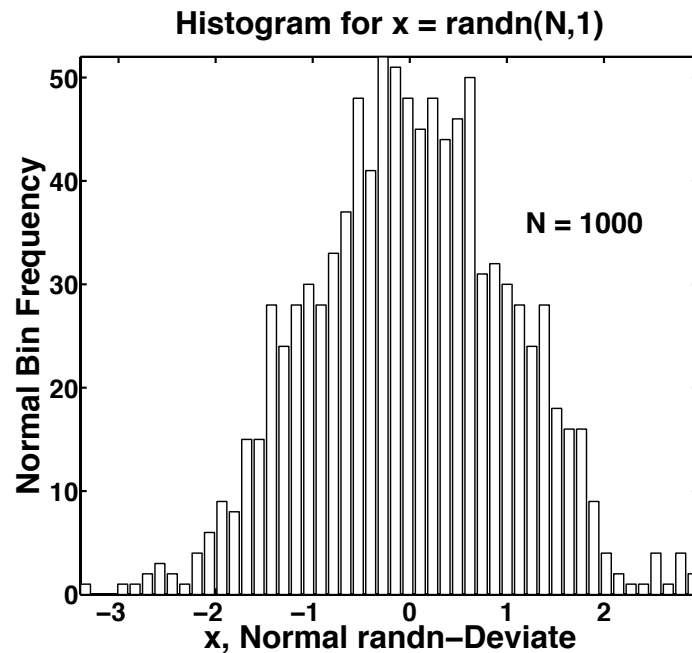


(b) Sample size $N = 10^5$.

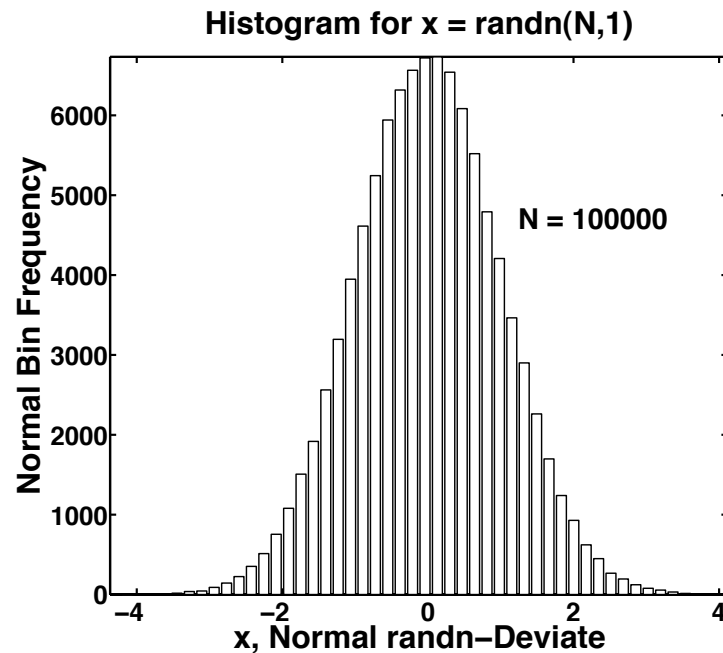
Figure 7: Histograms of simulations of the **standard uniform distribution** $f_X^{(u)}(x; 0, 1)$, $x \in (0, 1)$, using **MATLAB rand** for two different sample sizes N using *sample uniform code* in Hanson (2007) *Online Appendix B Preliminaries*, page B6. **Hence, need large sample size to get a reasonable approximation to the distribution.**

Uniform code example for Figure 7 abbreviated to one page:

```
function uniform09fig
% FINM 331 Uniform Histogram Demo. from Hanson book (2007)
fprintf('\nfunction uniform09fig OutPut:')
for m = 3:2:5;
    N=10^m;
    rand('twister',3);
    x=rand(N,1);
    xmean=mean(x); xstd=std(x);
    xmin = min(x); xmax = max(x);
    nbins = 30; % min(fix(sqrt(10^m)),101);
    xmin = 0; xmax = 1;
    xbin1 = xmin; xbin2 = xmax; dxbin = (xbin2-xbin1)/nbins;
    xbin = xbin1+dxbin/2:dxbin:xbin2-dxbin/2;
    nx = hist(x,xbin); % Need Edge Oriented histc.
    bar(xbin,nx); axis tight;
    title('Histogram for x = rand(N,1)', 'FontSize',44...
        , 'FontWeight', 'Bold');
% Code shorten for one page
end % For demo, use hist(rand(1e3,1),30), hist(rand(1e5,1),30)
% in the command window.
```



(a) Sample size $N = 10^3$.



(b) Sample size $N = 10^5$.

Figure 8: Histograms of simulations of the **standard normal distribution** with mean **0** and variance **1** using **MATLAB randn** with 50 bins for two sample sizes N . The histogram for the large sample size of $N = 10^5$ in Subfigure 8(b) exhibits a better approximation to the theoretical normal density $f_X^{(n)}(x; 0, 1)$ using *normal sample code* in Hanson (2007) *Online Appendix B Preliminaries*, page B9. Note that $\text{randn}(1e5, 1) \in (-5, 5)$, approximately, and $\text{normpdf}(5, 0, 1) = 1.4867e-6$.

Normal code example for Figure 8 abbreviated to one page:

```
function normal03fig1
% FINM 331 Normal Histogram Demo. from Hanson book (2007)
for m = 3:2:5
    N=10^m;
    x=randn(N,1);
    xmean=mean(x);
    xstd=std(x);
    xmin = min(x); xmax = max(x);
    xbin1 = xmin; xbin2 = xmax; dxbin = (xbin2-xbin1)/nbins;%
    xbin = xbin1+dxbin/2:dxbin:xbin2-dxbin/2;
    nx = hist(x,xbin); % Need Center Oriented hist.
    bar(xbin,nx); axis tight;
    title('Histogram for x = randn(N,1)', 'FontSize', 44);
    ks = [0.4,0.7]; nxmax = max(nx);
    xtext = xmax*ks(1);
    ytext=fix(ks(2)*nxmax);
    textn=['N = ' int2str(N)];
    text(xtext,ytext,textn,'FontSize',40,'FontWeight','Bold');
    ylabel('Normal Bin Frequency','FontSize',44);
    xlabel('x, Normal randn-Deviates','FontSize',44);
end% For demo., use hist(rand(1e3,1),30), hist(rand(1e5,1),30).
```


- *Being innovative can be helpful.* If the desired integral is of the form,

$$\int_a^b h(x)dx, \quad (59)$$

given a , b and a **moderately varying** $h(x)$, then let

$$f_X(x) = 1/(b - a) \quad (60)$$

be a *uniform density* on (a, b) ,

$$g(x) = (b - a) * h(x) \quad (61)$$

where $*$ denotes multiplication in **MATLAB**, forming a substitute expectation for the original integral (62),

$$\int_a^b g(x) f_X(x) dx = \mu_{g(X)} \equiv \mathbf{E}_X [g(X)]. \quad (62)$$

The sample mean is evaluated using the uniform variate,

$$X = a * \text{ones}(n, 1) + (b - a) * \text{rand}(n, 1); \quad (63)$$

an n -vector (**MATLAB is most efficient in matrix-vector form**).

The **sample mean** is

$$\text{gmean} = \text{mean}(g(X)); \quad (64)$$

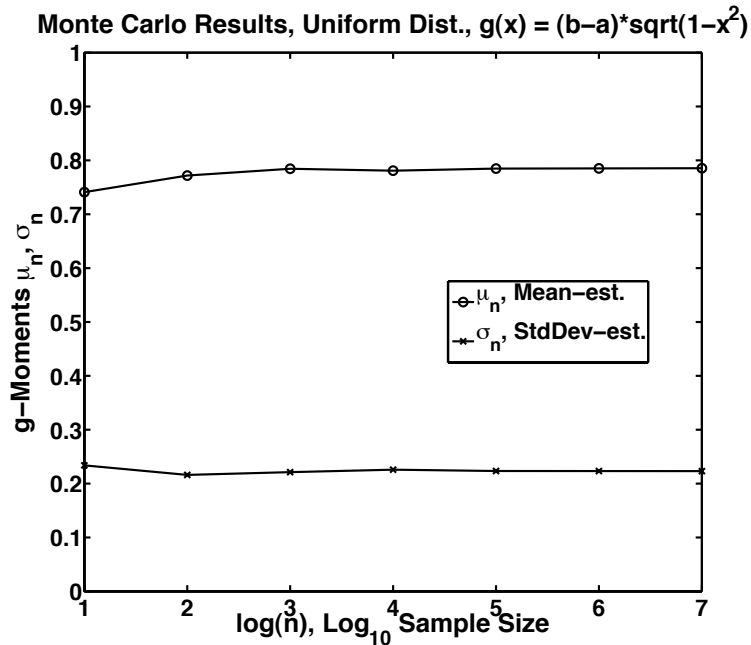
and the approximate integral is

$$\text{hintegral} = \text{mean}(g(X)) / (b - a); \quad (65)$$

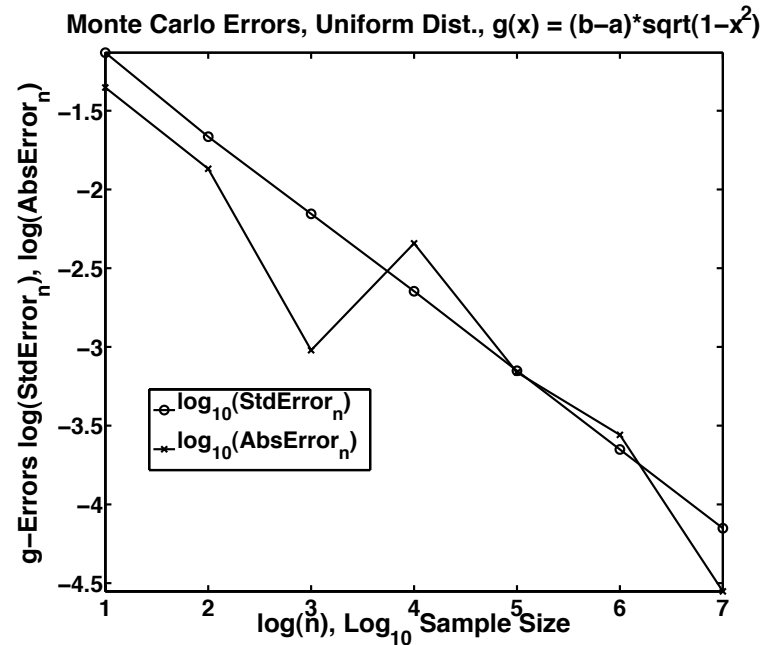
The **estimated standard error (SE)** of $\overline{g(X)}$ is

$$\sigma_{\overline{g(X)}} = \frac{\sigma_{g(X)}}{\sqrt{n}} \simeq \frac{\widehat{s}_n}{\sqrt{n}} = \widehat{\text{SE}}_n[\overline{g(X)}], \quad (66)$$

replacing the theoretical standard deviation $\sigma_{g(X)}$, which is usually unknown, by the feasibly estimated (unbiased as a variance) sample standard deviation \widehat{s}_n . The estimated standard error for the Monte Carlo approximation of the original will be $\widehat{\text{SE}}_n[\overline{g(X)}] / (b - a)$.



(a) Mean & Std. Dev.



(b) Std. & Abs. Errors.

Figure 9: Monte Carlo results for *uniform rand-deviates* for $f(x) = 1/(b-a)$ and $g(x) = (b-a) * h(x)$ on (a,b) as a function of sample size N using *sample MCM uniform code* in Hanson (2007) *applied stochastic text*, page 265.

- **Uniform Monte Carlo code example** for Figure 9 abbreviated to one page:

```
function mcmltest09
% FINM 331 mcmltest: Monte Carlo Method (2009),
% integral of h(x) = sqrt(1-x^2); on [a,b]:
a = 0; b = +1; % -1 <= a < b <= +1;
IntExact = 0.5*(asin(b)-asin(a))+0.5*(b*sqrt(1-b^2)-a*sqrt(1-a^2));
MugExact = IntExact;
Sigg = sqrt((b-a)^2*(1-(b^2+a*b+a^2)/3)-MugExact^2); kmax = 7;
n = zeros(1,kmax); meang = zeros(1,kmax); sigg = zeros(1,kmax);
sigdrn = zeros(1,kmax); error = zeros(1,kmax);
for k = 1:kmax
    rand('state',0); % set state or seed
    n(k) = 10^k; % sample size, k = log10(n(k)) ;
    x = a+(b-a)*rand(n(k),1); % get n(k) X 1 random sample on (a,b);
    g = (b-a)*sqrt(1-x.^2); % vectorized g;
    meang(k) = mean(g); % E[g(X)];
    sigg(k) = std(g); % sqrt(sigmag^2), sigmag^2 = unbiased Var(g );
    sigdrn(k) = sigg(k)/sqrt(n(k));
    error(k) = abs(meang(k)-MugExact);
    fprintf('%1i %8i %6.4f %6.4f %9.3e %9.3e %9.3e %9.3e\n'...
        ,k,n(k),meang(k),MugExact,sigg(k),Sigg,sigdrn(k),error(k))
end
% end mcmltest.m (For plot code see source)
```

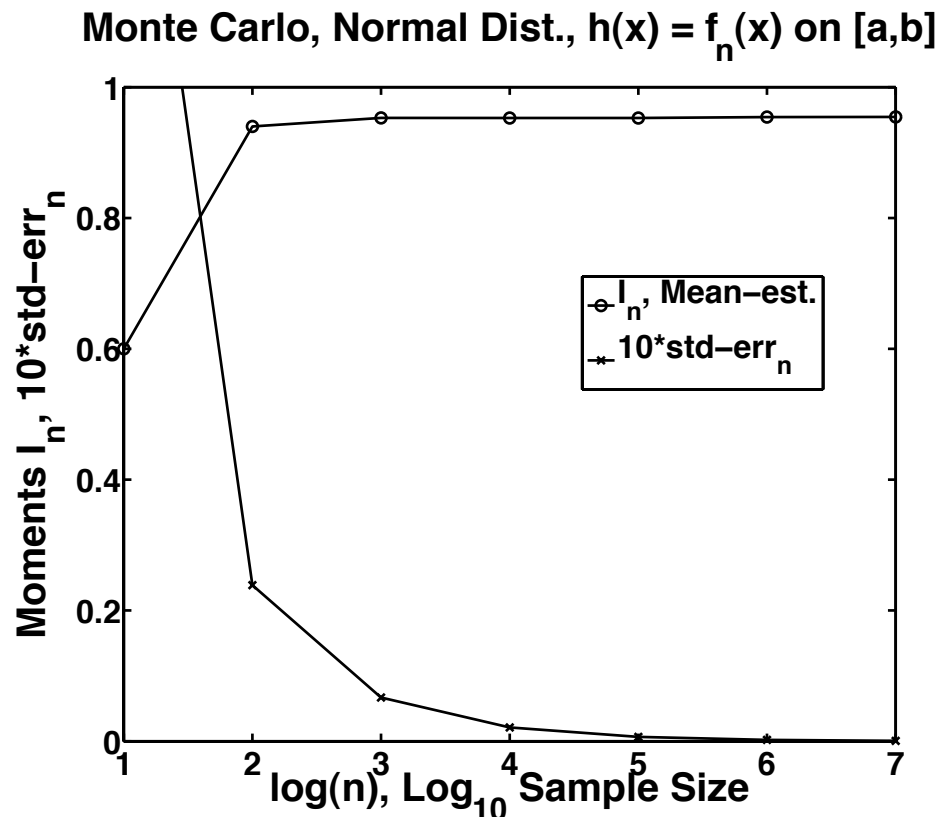


Figure 10: Monte Carlo results for *normal randn-deviates* for the standard normal distribution on (a, b) as a function of sample size N by the *von Neumann method of acceptance and rejection* to account for the reduced sample size on a finite interval (observations outside the interval are not counted) using *sample MCM uniform code* in Hanson (2007) *applied stochastic text*, page 272.

- *Normal Monte Carlo with Acceptance Sampling code example* for

Figure 10:

```
function mcm2acceptreject09
% FINM 331: Normal Monte Carlo with Acceptance-Rejection Sampling.
% I = int(h(x),x=a..b), h(x) = f(x) = exp(-x^2/2)/sqrt(2pi);
a = -2; b = 2; % limits of integration;
f = inline ('exp(-x.*x/2)./sqrt(2*pi)','x'); % x in [a,b]
kmax = 7; nmc = 10^kmax; nac = 0;% select MCM random sample size;
x = randn(1,nmc); % nmc X 1 normal distribution;
kv = zeros(1,kmax); Ihatn = zeros(1,kmax); stderrn = zeros(1,kmax);
for n = 1:nmc
    if (x(n) >= a) && (x(n) <= b)
        nac = nac + 1; % counts ACCEPTED points;
    end
    if n==10 || (n==100) || (n==1000) || (n==10000) || (n==100000) || (n==1000000) || (n==nmc)
        k = log10(n);
        kv(k) = k;
        Ihatn(k) = nac/n; % Estimate Integral
        stderrn(k) = sqrt(Ihatn(k)*(1-Ihatn(k))/(n-1));
        fprintf ('%2i   %8i   %8.6f   %9.3e\n',k,n,Ihatn(k),stderrn(k));
    end
end
% end mcm2acceptreject.m (For plot part of code see source.)
```

- **Acceptance-Rejection Sampling Method of von Neumann (1951):**

Often integrals will have a density $f_X(x)$ defined on $[a, b]$, but the integral is defined on a subset $[c, d] \subset [a, b]$ or in the more complicated case there is **NO random number generator** available for $f_X(x)$. In the simpler case, the **accepted (ac)** integral is

$$\begin{aligned} I^{(\text{ac})} &= \int_c^d g(x) f_X(x) dx = C^{(\text{ac})} \mathbf{E}[g(X) | X \in (c, d)] \\ &= C^{(\text{ac})} \mu_{g(X)}^{(\text{ac})}, \end{aligned} \quad (67)$$

where $C^{(\text{ac})} = \int_c^d f_X(x) dx$ is the **renormalization constant** and $\mu_{g(X)}^{(\text{ac})}$ is the mean of $g(X)$ on the accepted (ac) interval. Given the usual IID sample $\vec{X} = [X_i]_{n \times 1}$ on $[a, b]$ such that $n = n^{(\text{ac})} + n^{(\text{re})}$, i.e., $n^{(\text{ac})}$ is the **number of accepted observations**, $\{X_i^{(\text{ac})}\}$, in $[c, d]$ while $n^{(\text{re})}$ is the **rejected number**.

Also, let $\mathbf{1}^{(\text{ac})}(x) = 1$ if $x \in [c, d]$ and otherwise $\mathbf{0}$ be the **accepted indicator function**, so $g^{(\text{ac})}(x) \equiv g(x)\mathbf{1}^{(\text{ac})}(x)$ is the function on $[a, b]$, where we will work. The **unbiased sample approximation to the accepted normalization constant** is

$$C^{(\text{ac})} \simeq \frac{1}{n} \sum_{i=1}^n \mathbf{1}^{(\text{ac})}(X_i) = \frac{n^{(\text{ac})}}{n} \equiv \widehat{C}_n^{(\text{ac})}. \quad (68)$$

Then, the **accepted sample mean** of $g(X)$ is

$$\overline{g^{(\text{ac})}(X)}_n = \frac{1}{n} \sum_{i=1}^n g(X_i)\mathbf{1}^{(\text{ac})}(X_i) \equiv \widehat{I}_n^{(\text{ac})}, \quad (69)$$

since by expectations it is the **unbiased estimate** of

$$I^{(\text{ac})} = \mathbf{E}_X \left[\widehat{I}_n^{(\text{ac})} \right] = \mu_{\widehat{I}_n^{(\text{ac})}}, \text{ as well as } \mu_{g^{(\text{ac})}(X)}.$$

Similarly, the variance is by the usual IID calculation,

$$\begin{aligned}\sigma_{\hat{I}_n^{(\text{ac})}}^2 &= \text{Var}_X \left[\hat{I}_n^{(\text{ac})} \right] = \mathbf{E}_X \left[\left(\hat{I}_n^{(\text{ac})} - \mu_{\hat{I}_n^{(\text{ac})}} \right)^2 \right] \\ &\stackrel{\text{iid}}{=} \frac{1}{n} \mathbf{E}_X \left[\left(g^{(\text{ac})}(X_i) - \mu_{g^{(\text{ac})}(X)} \right)^2 \right] = \frac{\sigma_{g^{(\text{ac})}(X)}^2}{n},\end{aligned}\tag{70}$$

The **unbiased sample variance** of $g^{(\text{ac})}(X)$ is

$$\hat{s}_n^2 [g^{(\text{ac})}] = \frac{1}{n-1} \sum_{i=1}^n \left(g^{(\text{ac})}(X_i) - \overline{g^{(\text{ac})}(X)}_n \right)^2\tag{71}$$

and thus the **standard error** of $g(X)$ can be approximated by

$$\text{SE} \left[\hat{I}_n^{(\text{ac})} \right] = \frac{\sigma_{g^{(\text{ac})}(X)}}{\sqrt{n}} \simeq \frac{\hat{s}_n [g^{(\text{ac})}]}{\sqrt{n}}.\tag{72}$$

In Figure 10, $g(x) = 1$, so

$$\text{SE}_n^{(\text{ac})} \left[\hat{I}_n^{(\text{ac})} \right] \simeq \sqrt{\frac{1}{n} \frac{n^{(\text{ac})}}{n} \left(1 - \frac{n^{(\text{ac})}}{n} \right)}.\tag{73}$$

Remarks: The details are left to the viewer. Also, details of general acceptance-rejection techniques will be given later.

- **Monte Carlo Risk-Neutral European Call Pricing code example:**

Note automatic acceptance and rejection by max function.

```
function mcm4eurocall
% Monte Carlo Risk-Neutral Pricing for European Call Options (2009),
% Adapted from D.J. Higham, Comp.Sci.&Engr., Nov/Dec2004, pp.72-79.
clc; clear
fprintf('Monte Carlo Pricing for European Call Options (2009)');
A0 = 2; K = 1; r = 0.05; sigma = 0.25; T = 3; N = 1.e6; PC = 95/100;
randn('state',50);
A = A0*exp((r-0.5*sigma^2)*T+sigma*sqrt(T)*randn(N,1)); %AssetPrice
C = exp(-r*T)*max(A-K,0); % Call Prices with Acceptance by Max;
Cmean = mean(C);
SE_C = std(C)/sqrt(N); % Standard Error for Call Cmean
width = norminv((1+PC)/2)*SE_C; % 100*PC% CI Bandwidth
CImin = Cmean-width; % Confidence Inteval left
CImax = Cmean+width; % Confidence Inteval right
fprintf('\nCImin=%6.4f < Cmean=%6.4f < CImax=%6.4f;\n' ...
,CImin,Cmean,CImax);
%%%%%%%% Delete this function if have Statistics Toolbox %%%
function z = norminv(p)
z = -sqrt(2)*erfcinv(2*p);
% end mcm3eurocall.m
```

Output: CImin=1.1436 < Cmean=1.1453 < CImax=1.1471;

Remark: Cmean is close to Black-Scholes answer, $C_{bs} = 1.1447$, and within CI.

*Lecture 1 Homework Posted in Chalk Assignments,
due by Lecture 2 in Chalk Assignments!*

Summary of Lecture 1:

- 1. Review Syllabus**
- 2. Additive vs. Multiplicative Models Examples**
- 3. Sampling and Independent Observation Statistics**
- 4. Law of Large Numbers (LLN)**
- 5. Central Limit Theorem (CLT)**
- 6. Monte Carlo**