$Gabe \in \mathbb{R}$
GabeL@virginia.edu
https://gabriel-vzh2vs.github.io/SYS3062-Website/
UVA's Systems and Information Engineering

January 27th, 2026

SYS 3062: Lab 2

Skewness, Kurtosis, and Retiring with Monte Carlo

Beyond the Average: Why Skewness and Kurtosis Matter

In Monte Carlo simulations, the Mean (μ) and Standard Deviation (σ) only tell part of the story. To understand risk, we need to look at the shape of the data.

Skewness (Asymmetry)

- Measures if the data is biased to the left or right.
- **Why it matters:** In retirement, we care if the uncertainty leans towards "making extra money" (positive skew) or "losing principal" (negative skew).

Kurtosis (The Tails)

- Measures the "heaviness" of the tails (frequency of outliers).
- **Why it matters:** High kurtosis warns us of *extreme events*.

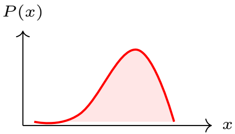
GabeL@virginia.edu

What is Skewness

Definition

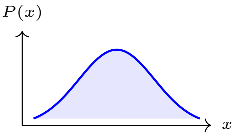
Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean.

Negative Skew
(Skew < 0)



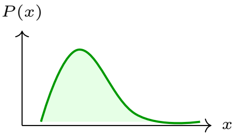
"Tail is on the Left"

Symmetric
(Skew $= 0$)



"Normal Distribution"

Positive Skew
(Skew > 0)



"Tail is on the Right"

GabeL@virginia.edu

Calculating Skewness (Fisher's)

Fisher's Definition (γ_1)

The moment coefficient of skewness of a random variable X is the third standardized moment:

$$\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}}$$

Where:

- $\mu = E[X]$ is the mean.
- σ is the standard deviation.
- $E[\cdot]$ is the expectation operator.
- μ_3 is the third central moment.
- Sometimes denoted using cumulants: $\frac{\kappa_3}{\kappa_2^{3/2}}$.

Expanding the Calculation

If μ and σ are finite, we can express skewness using non-central moments $E[X^3]$.

Derivation

$$\begin{aligned}
\gamma_1 &= E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \\
&= \frac{E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] - \mu^3}{\sigma^3} \\
&= \frac{E[X^3] - 3\mu(E[X^2] - \mu E[X]) - \mu^3}{\sigma^3} \\
&= \frac{E[X^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3}
\end{aligned}$$

This formula allows us to calculate skewness directly from the raw moments of the data.

GabeL@virginia.edu

Skewness of a Sample (b_1)

When dealing with a **sample** of data, we estimate skewness using the following formula:

Sample Skewness Formula

$$b_1 = \frac{m_3}{s^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}}$$

Component Definitions:

- n: Sample size.
- $\bar{x}$: Sample mean.
- m_3: The (biased) sample third central moment (Numerator).
- s: The sample standard deviation (derived from the sample variance in the Denominator).

Note: The denominator represents the sample variance raised to the power of $3/2$ (which is the standard deviation cubed).

Dataset: $X = \{1, 1, 4\} \qquad (n = 3)$

Step 1: Mean ($\bar{x}$)

$$\bar{x} = \frac{1+1+4}{3} = 2$$

Step 2: Sample Variance Term (Denominator part)

$$s^2 = \frac{1}{3-1}\left[(1-2)^2 + (1-2)^2 + (4-2)^2\right] = \frac{1}{2}[1+1+4] = 3$$

$$\text{Denominator } (s^3) = (3)^{3/2} \approx 5.196$$

Step 3: Third Moment Term (m_3)

$$m_3 = \frac{1}{3}\left[(1-2)^3 + (1-2)^3 + (4-2)^3\right] = \frac{1}{3}[-1-1+8] = \frac{6}{3} = 2$$

Step 4: Final Skewness (b_1)

$$b_1 = \frac{2}{5.196} \approx \mathbf{0.38} \quad \text{(Positively Skewed)}$$

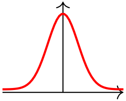
GabeL@virginia.edu

Kurtosis: Definition & Shapes

What is Kurtosis?

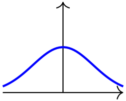
Kurtosis is a statistical measure that describes the shape of the tails of a probability distribution. It quantifies how often extreme values (outliers) occur compared to a normal distribution.

Leptokurtic



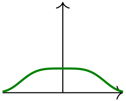
Heavy tails, prone to outliers.

Mesokurtic



Baseline (Gaussian).

Platykurtic



Light tails, fewer outliers.

GabeL@virginia.edu

Calculating "Good Enough" Kurtosis

Population Kurtosis The fourth standardized moment:

$$\text{Kurt}[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{\mu_4}{\sigma^4}$$

Biased Sample Estimator (g_2) For a sample, we often use a simpler, but **biased**, estimator:

Biased Estimator Formula

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n}\sum(x_i-\bar{x})^4}{\left[\frac{1}{n}\sum(x_i-\bar{x})^2\right]^2} - 3$$

Note on Bias: An estimator is "biased" if its expected value differs from the true population parameter. This formula tends to underestimate kurtosis in small samples.

Unbiased Kurtosis (A Common Formula)

To correct for bias (especially in small sample sizes), we use a significantly more complex formula. This is the version used by software like Excel, SPSS, and SAS.

Unbiased Estimator (G_2)

$$G_2 = \underbrace{\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^4}_{\text{Shape Adjustment}} - \underbrace{\frac{3(n-1)^2}{(n-2)(n-3)}}_{\text{Baseline Correction}}$$

Why is it so complex?

- It accounts for the degrees of freedom lost by estimating the mean ($\bar{x}$) and standard deviation (s).
- It ensures that if the population is Normal, the expected value is exactly 0.

Why Learn Kurtosis and Skewness?

The Trap of Normality: In simulation, relying solely on Mean and Standard Deviation (assuming a Normal distribution) can hide critical risks.

Kurtosis: Assessing Extreme Risk

Focus: The "Tails" (Outliers).

- **Simulation Insight:** High kurtosis warns of "Black Swan" events, rare but catastrophic outcomes.
- **Key Applications:**
  - Financial Modeling (Market Crashes)
  - Weather Forecasting (Extreme Storms)
  - Network Traffic (Data Bursts)

Skewness: Identifying Bias

Focus: Asymmetry.

- **Simulation Insight:** Outcomes are naturally biased in one direction. Symmetric confidence intervals will fail here.
- **Key Applications:**
  - Queuing Systems
  - Risk Assessment (Losses vs. Gains)
  - Reliability Engineering

GabeL@virginia.edu

The Real-World Problem

Standard financial advice often uses simple, deterministic formulas ($I = Prt$) to project retirement savings.

- **The Flaw:** This assumes returns are constant every single year.
- **The Reality:** Markets crash, interest rates fluctuate, and "average" returns can hide dangerous risks.

Lab Objectives:

1. **Compare Models:** Contrast a static model against a stochastic (random) Monte Carlo model.
2. **Analyze Extreme Events:** Use simulation to find the likelihood of "tail risks" (e.g., running out of money) rather than just finding the mean ending balance.

GabeL@virginia.edu

Lab 2: Objectives & Concepts

Primary Purpose

To demonstrate how Monte Carlo simulation improves upon deterministic models for retirement planning by introducing uncertainty and analyzing extreme events.

The Three Sub-models:

1. **Sub-model 1 (Deterministic):** A fixed growth formula ($I = Prt$ style) using constant rates.
2. **Sub-model 2 (Single Stochastic):** Introduces randomness to the interest rate (r) for a single run.
3. **Sub-model 3 (Full Monte Carlo):**
   - Randomizes **Principal** (P_0) per experiment.
   - Randomizes **Interest Rate** (r) per year.
   - Runs N experiments to analyze **Skewness** and **Kurtosis**.

Algorithm Design: Sub-model 3

Algorithm 1 Monte Carlo Retirement Simulation

$Results \leftarrow []$

$N \leftarrow 1000$ ▷ Number of experiments

for $i = 1$ **to** N **do**

  $P \leftarrow$ Triangular$(15k, 25k, 35k)$ ▷ Sample Principal once

  for $t = 1$ **to** 40 **do**

    $r \leftarrow$ Normal$(\mu = 0.10, \sigma = 0.12)$ ▷ Sample Rate annually

    $P \leftarrow P \times (1 + r) + 10,000$ ▷ Apply Interest + Contribution

  end for

  Append P to $Results$

end for

Calculate Skewness$(Results)$

Calculate Kurtosis$(Results)$

GabeL@virginia.edu

The code maps directly to the algorithm. Note the separation of the **Experiment Loop** and the **Yearly Projection Loop**.

```
def run_pattern_3(params):
    n = params['n_experiments']
    t = params['years_t']
    s = params['contribution_s']
    final_values = []

    # 1. Experiment Loop
    for i in range(n):
        # Sample Stochastic Principal ONCE per experiment
        p_current = np.random.triangular(
            params['p_tri_min'], params['p_tri_mode'], params['p_tri_max'])

        # 2. Yearly Projection Loop
        for _ in range(t):
            # Sample Stochastic Rate EACH year
            r = np.random.normal(params['mean_r'], params['sigma_r'])
            p_current = p_current * (1 + r) + s

        final_values.append(p_current)
    return np.array(final_values)
```

Going Further: Sensitivity Analysis

How does volatility (σ) affect the shape of our outcomes? We can sweep through sigma values to see how Skewness and Kurtosis react.

Code Concept:

- Wrap the experiment in a loop over ‘sigma-range‘.
- Collect Mean, Skew, and Kurtosis for each level of volatility.

Why it matters:

- High Kurtosis = Higher risk of extreme poverty or extreme wealth.
- Helps clients understand "Tail Risk" beyond just the "Average Return."

Key Output Metrics

Use `scipy.stats.skew` and `scipy.stats.kurtosis` on the final array of results to quantify these risks.

Now, it's your turn to implement a retirement plan in Python or using XLRisk! Good Luck!

How to Naively Pick a Distribution

Discrete Data (Counts/Integers)

Can you estimate outcomes & probs?

- **Yes:** Use an **Empirical Distribution**.
- **No:** Check Symmetry.

Is the data Symmetric?

- **Yes:**
  - Clustered around center? → **Binomial**
  - Equal probability? → **Uniform Discrete**
- **No:**
  - Outliers positive? → **Geometric, Negative Binomial**

Continuous Data (Measurements)

Is the data Symmetric?

- **Yes:**
  - Clustered around center? → **Normal, Logistic**
  - No central cluster? → **Uniform, Triangular**

Is the data Asymmetric?

- **Where are the outliers?**
  - Only Positive? → **Exponential, Lognormal, Gamma, Weibull**
  - Mostly Negative? → **Min. Extreme Value**

GabeL@virginia.edu

Feedback Form

Throughout the course, I may ask for your feedback on:

- **Course Materials**
- **Lecture & Lab**
- **Assignments**

GabeL@virginia.edu

Action Items

First, I will thank you all again for coming to lab at 5 pm on a Tuesday!

And now, you need to do the following tasks:

- **Submit** your Python file to Gradescope by next week's lab!
- **Read** Prelab 3 for next week!

Let's keep on moving forward!

GabeL@virginia.edu