

Sharpe Ratio Practice Note

Geng Deng, PhD, FRM* Tim Dulaney, PhD[†]

Craig McCann, PhD, CFA[‡]

Securities Litigation and Consulting Group, Inc.

3998 Fair Ridge Drive, Suite 250, Fairfax, VA 22033

June 26, 2012

1 Introduction

Following the seminal work of Roy (1952), Sharpe (1966) suggested a simple metric to compare investment return distributions.¹ The Sharpe ratio (SR) is defined to be the ratio of the average expected excess return over the risk-free rate (μ) to the standard deviation of the expected excess returns (σ):

$$SR = \frac{\mu}{\sigma}.$$

Higher Sharpe ratios indicate more compensation for each unit of investment risk. Of course, μ and σ are unobservable and must be estimated using historical data.² Given a sample of historical returns $\{R_1, R_2, \dots, R_n\}$ and a constant risk-free rate R_f , the statistical estimator for the Sharpe ratio is given by

$$\widehat{SR} = \frac{\hat{\mu}}{\hat{\sigma}}$$

where

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n (R_i - R_f) \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (R_i - R_f - \hat{\mu})^2.$$

*Director of Research, Office: (703) 539-6764, Email: <mailto:GengDeng@slcg.com>.

[†]Senior Financial Economist, Office: (703) 539-6777, Email: <mailto:TimDulaney@slcg.com>.

[‡]President, Office: (703) 539-6760, Email: <mailto:CraigMcCann@slcg.com>.

¹For a review of the definitions of the moments of a distribution, see A.

²“Expected excess return” is a bit of a misnomer in the context of historical data. Here we simply mean the average excess return expected based upon the series of historical returns considered within the sample.

We provide a detailed discussion of the relationship between the underlying return distribution and the estimated Sharpe ratio in the next section, but it is important to note the distinction between the ex-post Sharpe ratio estimated from historical returns and the Sharpe ratio of the underlying return distribution (which is not observable). If the riskless rate is not constant, then the appropriate riskless rate in each period should be used to determine the excess returns. The mean and variance of the sample should then be computed with respect to these excess returns.

The usefulness of the Sharpe ratio is most easily shown through example. Consider two investments with normal return distributions with mean annual excess return given by $\hat{\mu}_1 = 12\%$ and $\hat{\mu}_2 = 15\%$, and standard deviation of excess returns given by $\hat{\sigma}_1 = 20\%$ and $\hat{\sigma}_2 = 30\%$. Although the second investment has higher excess returns on average, the volatility of these returns is also higher. Computing the Sharpe ratio for these two investments gives: $\widehat{SR}_1 = 0.6$ and $\widehat{SR}_2 = 0.5$. Although the first investment has smaller excess returns on average, the first investment compensates investors more for each unit of investment risk.³

The Sharpe ratio assumes that asset returns are independent and identically distributed (IID) normal variables; however, many assets do not exhibit such distributions. For example, the return distribution of hedge funds are often negatively skewed and exhibit positive excess kurtosis. Lo (2002) showed that the positive serial correlation of hedge fund returns leads to an overstatement of annualized Sharpe ratios.⁴

2 Sharpe Ratio Distributions

2.1 Independent and Identically Distributed Normal Returns

Independent and identically distributed returns is a common assumption and a simple place to start with the analysis of the statistical distribution of Sharpe ratios. Assuming that an investment returns $\{R_1, R_2, \dots, R_n\}$ are independent and identically distributed with finite mean μ and variance σ^2 , the following relations in the limit

$$\begin{aligned}\sqrt{n}(\hat{\mu} - \mu) &\Rightarrow N(0, \sigma^2) \\ \sqrt{n}(\hat{\sigma}^2 - \sigma^2) &\Rightarrow N(0, 2\sigma^4)\end{aligned}$$

as a result of the Central Limit Theorem.⁵

These equations imply that the variance in the estimators $\hat{\mu}$ and $\hat{\sigma}^2$ take the following asymptotic forms,

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n} \quad \text{and} \quad \text{Var}(\hat{\sigma}^2) = \frac{2\sigma^4}{n}$$

³A careful reader might ask what the probability is that the first investment's return distribution actually possesses a higher Sharpe ratio than the second and at what significance. We will address such questions later in the text.

⁴Serial correlation means that the return for the current period is correlated with the return for the previous period.

⁵ $N(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2 .

and therefore the sampling error from these estimators decreases with increasing sample size. Using Taylor's theorem, one can show that if $g = g(\mu, \sigma^2)$, then

$$\sqrt{n} (g(\hat{\mu}, \hat{\sigma}^2) - g(\mu, \sigma^2)) \Rightarrow N(0, V_{IID})$$

where

$$V_{IID} = \sigma^2 \left(\frac{\partial g}{\partial \mu} \right)^2 + 2\sigma^4 \left(\frac{\partial g}{\partial \sigma^2} \right)^2.$$

In particular, if $g(\mu, \sigma^2) = \mu/\sigma$, then

$$V_{IID} = 1 + \frac{1}{2} \left(\frac{\mu}{\sigma} \right)^2.$$

As a result, the standard deviation in the estimated Sharpe ratio is then given by

$$\sigma(\widehat{SR}) \Rightarrow \sqrt{\frac{1}{n} \left(1 + \frac{1}{2} SR^2 \right)} \approx \sqrt{\frac{1}{n-1} \left(1 + \frac{1}{2} \widehat{SR}^2 \right)} = \hat{\sigma}(\widehat{SR}).$$

For example, consider a situation wherein a Sharpe ratio estimated from 10 years of monthly returns ($n = 120$) is one ($\widehat{SR} = 1$). At a 95% confidence level, one knows that the population mean of the Sharpe ratio is within 0.219 of 1. In other words, $0.781 \leq SR \leq 1.219$ at a 95% confidence level.

For smaller Sharpe ratios ($SR < 1$), the majority of the Sharpe ratio estimation error is due to the uncertainty in the population mean (μ). For larger Sharpe ratios ($SR > 2$), the majority of the Sharpe ratio estimation error is due to the uncertainty in the population variance (σ^2).

2.2 Non-normal IID Returns

Mertens (2002) shows how the results of the previous section can be generalized to non-normal return distributions. He considers returns $\{R_1, R_2, \dots, R_n\}$ drawn from a distribution with finite mean μ , variance σ^2 , skewness γ_3 and kurtosis γ_4 .⁶

In particular, we have the following alteration of the distribution of Sharpe ratio estimators

$$\sqrt{n} (g(\hat{\mu}, \hat{\sigma}^2) - g(\mu, \sigma^2)) \Rightarrow N(0, V)$$

where

$$V = 1 + \frac{1}{2} \left(\frac{\mu}{\sigma} \right)^2 - \left(\frac{\mu}{\sigma} \right) \hat{\gamma}_3 + \left(\frac{\mu}{\sigma} \right)^2 \left(\frac{\hat{\gamma}_4 - 3}{4} \right).$$

⁶Opdyke (2007) showed that his expressions are actually applicable under more general assumptions of stationary and ergodic returns distributions which were independently derived by Christie (2005).

The standard deviation of the estimated Sharpe ratio is then given by⁷

$$\hat{\sigma}(\widehat{SR}) = \sqrt{\frac{1}{n-1} \left(1 + \frac{1}{2} \widehat{SR}^2 - \widehat{SR} \hat{\gamma}_3 + \widehat{SR}^2 \left(\frac{\hat{\gamma}_4 - 3}{4} \right) \right)}.$$

For return distributions that exhibit negative skewness ($\hat{\gamma}_3 < 0$) and positive excess kurtosis ($\hat{\gamma}_4 > 3$), the volatility in the estimate of the Sharpe ratio may be much higher than that of a similar Sharpe ratio assuming a normal return distribution with similar mean and variance.

3 Probabilistic Sharpe Ratio (PSR) and Minimum Track Record (MTR)

3.1 Definitions

When comparing two Sharpe ratios, one must take into account the estimation errors involved with each computation. Bailey and de Prado (2011) suggest a measure for the comparison of an estimated Sharpe ratio (\widehat{SR}) to a benchmark Sharpe ratio (SR^*). In particular, they define the probabilistic Sharpe ratio as the probability that the observed Sharpe ratio is larger than the benchmark Sharpe ratio. More explicitly,

$$\widehat{PSR}(SR^*) = \text{Prob}(\widehat{SR} > SR^*) = 1 - \int_{-\infty}^{SR^*} \text{Prob}(\widehat{SR}) d\widehat{SR}.$$

Applying the results derived above assuming the return distribution is non-normal, we have

$$\widehat{PSR}(SR^*) = \Phi \left[\frac{(\widehat{SR} - SR^*) \sqrt{n-1}}{\sqrt{1 + \frac{1}{2} \widehat{SR}^2 - \widehat{SR} \hat{\gamma}_3 + \widehat{SR}^2 \left(\frac{\hat{\gamma}_4 - 3}{4} \right)}} \right]$$

where Φ is the cumulative distribution function for the standard normal distribution. A $\widehat{PSR}(SR^*) > 0.95$ indicates that the estimated Sharpe ratio is greater than the benchmark Sharpe ratio at a 95% confidence level.

This probabilistic measure used for the comparison of Sharpe ratios is dependent upon the sample size (n). All else being equal, we require a larger sample size for a return distribution that exhibits negative skewness and positive excess kurtosis when compared to a normal distribution to have the same confidence level.

⁷Although the underlying return distribution is not normal, the distribution of Sharpe ratio estimators follows a normal distribution.

Bailey and de Prado (2011) give a measure for the minimum sample size required to accept the proposition that $\widehat{SR} > SR^*$ at a significance of $\alpha\%$. They define the minimum track record (MTR) as

$$\text{MTR} = 1 + \left[1 + \frac{1}{2}\widehat{SR}^2 - \widehat{SR}\widehat{\gamma}_3 + \widehat{SR}^2 \left(\frac{\widehat{\gamma}_4 - 3}{4} \right) \right] \left(\frac{Z_\alpha}{\widehat{SR} - SR^*} \right)^2.$$

As the skewness of the return decreases, or the excess kurtosis increases, the minimum track record increases.

3.2 Example: Minimum Track Record

The MTR allows an analyst to determine the number of observations of a return distribution that are required to conclude that the Sharpe ratio estimator is larger than a benchmark at a prespecified confidence level. As an example, assume that the information in Table 1 describes the distribution of excess monthly returns of a particular investment. Assume further that we would like to determine the minimum number of observations required to conclude that the observed Sharpe ratio is larger than the benchmark Sharpe ratio $SR^* = 0$ at a 95% confidence level ($Z_\alpha = 1.96$).

Table 1: Asset parameters

Parameter	Asset One
Expected Excess Return	0.62%
Standard Deviation	1.00%
Sharpe Ratio (\widehat{SR})	0.444
Skewness	-1.2
Kurtosis	9.4
Sharpe Ratio Standard Error	12.8%

Assuming that the parameters of the historical return distribution (in Table 1) remain relatively constant across our sample size, we would have to make at least 39 monthly observations to be confident that the Sharpe ratio of the underlying return distribution was significantly different from the benchmark $SR^* = 0$ at the 95% confidence level.

4 Sharpe Ratio Efficient Frontier

Bailey and de Prado (2011) define a Sharpe Ratio Efficient Frontier (SEF) by taking analogy with the original mean-variance framework of the seminal work of Markowitz (1952). The authors define the SEF as the set of portfolios that deliver the highest expected excess return on risk subject to

the level of uncertainty surrounding those portfolios' excess return on risk. More explicitly, for a given level of uncertainty σ^* , the Sharpe ratio efficient frontier is given by

$$\text{SEF}(\sigma^*) = \left\{ P \mid \max_{\hat{\sigma}_{\widehat{SR}(P)} = \sigma^*} \widehat{SR}(P) \right\}.$$

This approach deals with uncertainty in the model's input parameters directly. The usual mean-variance/efficient frontier framework's sensitivity to input estimates is recognized and incorporated into the SEF developed by Bailey and de Prado (2011).

In order to understand the usefulness of this SEF in portfolio selection, it is instructive to study a few examples. We begin with a simple example that involves two uncorrelated assets and move on to consider an example with three uncorrelated assets.

4.1 Example One: Two Uncorrelated Assets

For the first example, consider two investments with return distributions as listed in Table 2. Assume that a decade of monthly returns have been observed for each investment ($n = 120$).

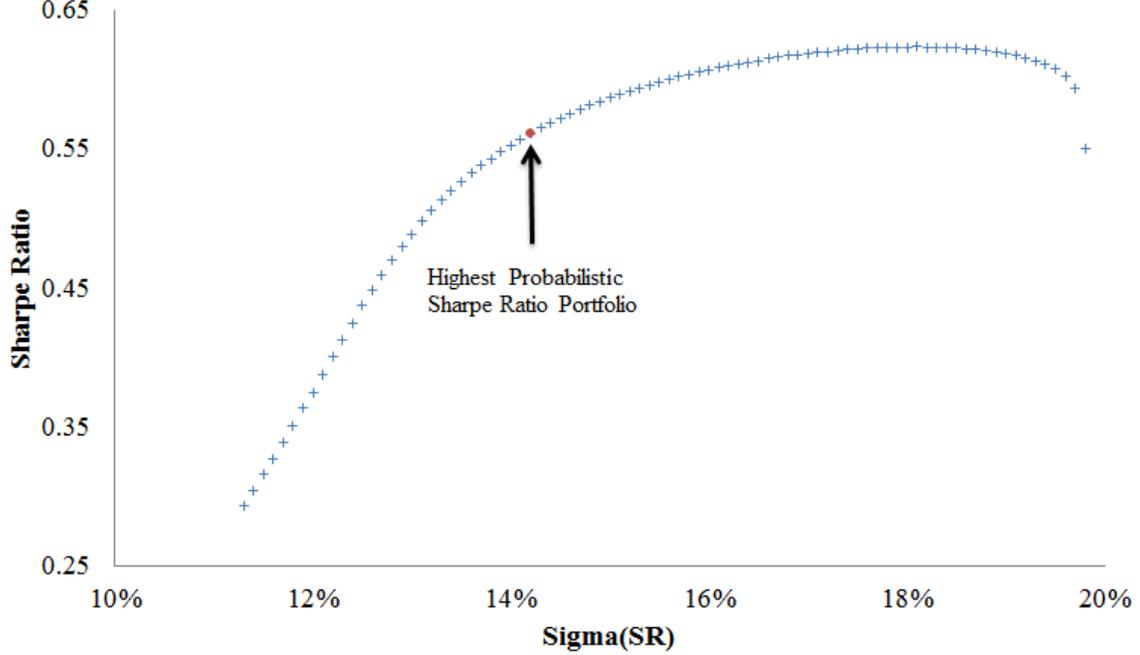
Table 2: Monthly return distribution parameters for a decade of monthly returns

Parameter	Asset One	Asset Two
Expected Excess Return	0.41%	1.27%
Standard Deviation	1.40%	2.31%
Sharpe Ratio (\widehat{SR})	0.293	0.550
Skewness	-1.2	-5.8
Kurtosis	9.4	5.2
Sharpe Ratio Standard Error	11.3%	19.5%

In Figure 1 we have plotted the Sharpe ratio efficient frontier for portfolios that invest only in asset one and asset two as described in Table 2. We have considered the set of portfolios with weight $w_1 = 1 - i/1000$ in asset one and weight $w_2 = i/1000$ in asset two ($i \in \{0, 1, 2, \dots, 1000\}$). We have assumed that the two assets are uncorrelated.

A portfolio that invests only in asset two is not efficient since the Sharpe ratio for such a portfolio is 0.550 and this is smaller than that of another portfolio with weights $\{w_1 = 30.7\%, w_2 = 69.3\%\}$ and Sharpe ratio 0.607 (for the same Sharpe ratio standard error, 19.5%). The portfolio that delivers the highest probabilistic Sharpe ratio is that which maximizes the quantity $\widehat{SR}/\hat{\sigma}_{\widehat{SR}}$. In this case, this portfolio has weights $\{w_1 = 69.3\%, w_2 = 30.7\%\}$, Sharpe ratio 0.561 and Sharpe ratio uncertainty 14.2%. This portfolio is pointed out in Figure 1.

Figure 1: Sharpe ratio efficient frontier for a portfolio consisting of an investment in asset one and asset two; The portfolio that delivers the highest probabilistic Sharpe ratio (indicated with a black arrow)



4.2 Example Two: Three Uncorrelated Assets

For the second example, consider three investments with return distributions as listed in Table 3. Assume once again that a decade of monthly returns have been observed for each investment ($n = 120$).

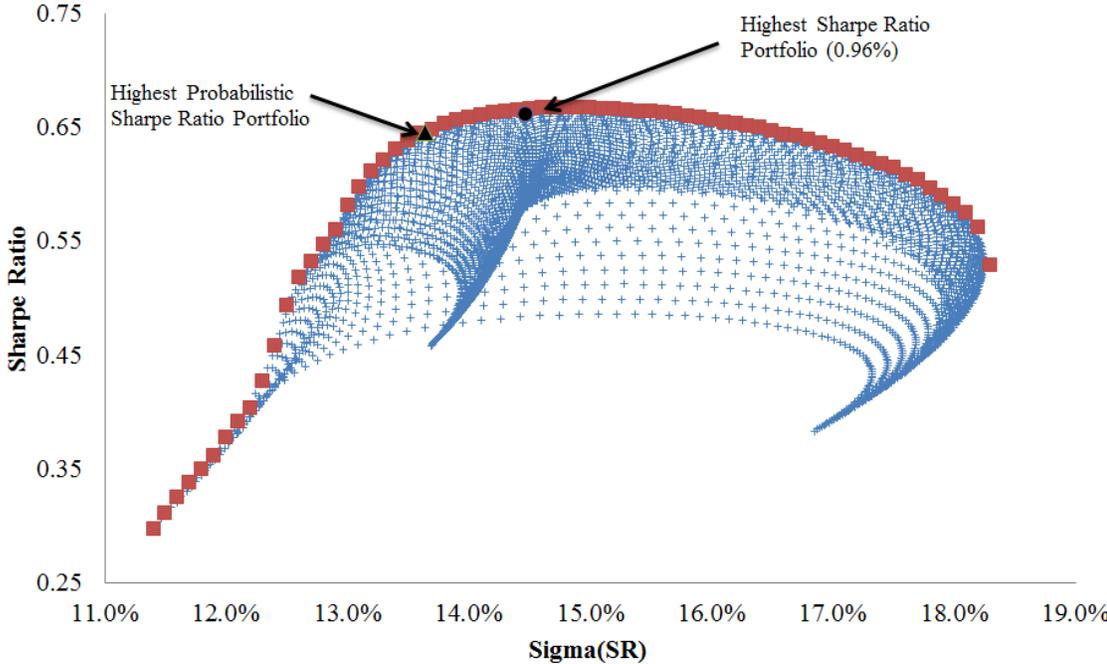
Table 3: Monthly return distribution parameters for a decade of monthly returns

Parameter	Asset One	Asset Two	Asset Three
Expected Excess Return	0.29%	0.86%	1.33%
Standard Deviation	0.97%	2.24%	2.90%
Sharpe Ratio (\widehat{SR})	0.299	0.384	0.459
Skewness	-1.2	-5.8	-2.4
Kurtosis	9.4	5.2	3.5
Sharpe Ratio Standard Error	11.4%	16.9%	13.7%

There are a total of 5,151 unique portfolios with weights $\{w_1 = i/100, w_2 = j/100, w_3 = 1 - (i + j)/100\}$ where $\{i, j\} \in \{0, 1, 2, \dots, 100\}$. In Figure 2 we plot the Sharpe ratio for each

portfolio as a function of the Sharpe ratio standard error. We assume that the three assets are uncorrelated.

Figure 2: Sharpe ratios of portfolios investing in assets one, two and three as a function of Sharpe ratio standard error (blue pluses); Sharpe ratio efficient frontier (red squares)



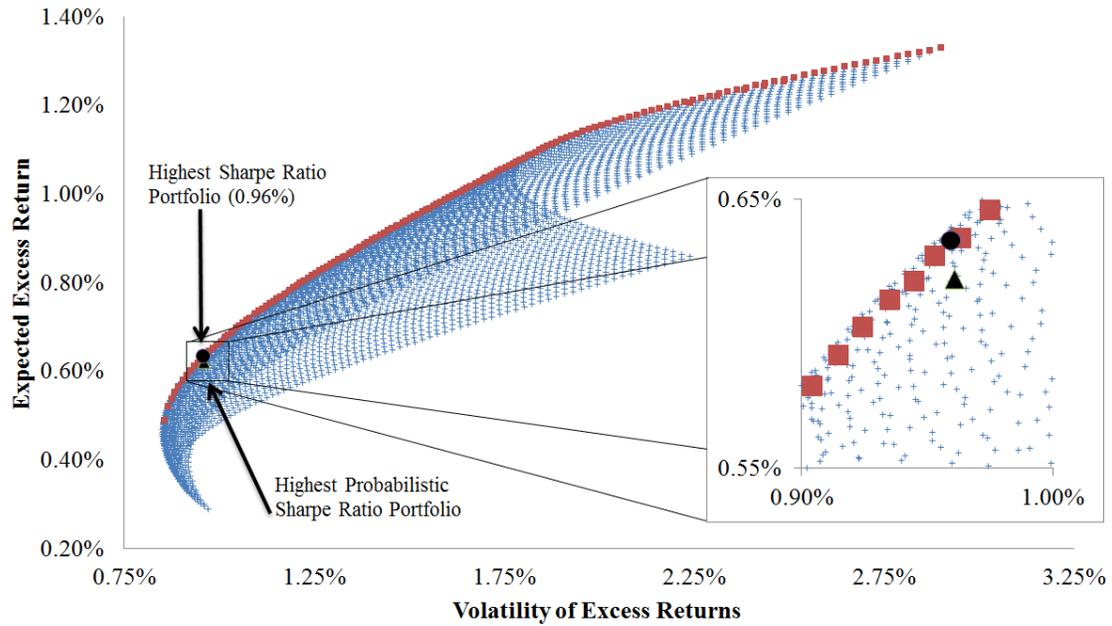
A portfolio that invests only in asset two or asset three is not efficient in this case. Rather than investing in asset two alone ($\widehat{SR} = 0.384$ and $\hat{\sigma}_{\widehat{SR}} = 16.9\%$), one could invest with portfolio weights $\{w_1 = 36\%, w_2 = 42\%, w_3 = 22\%\}$ realizing $\widehat{SR} = 0.638$ with the same Sharpe ratio standard error ($\hat{\sigma}_{\widehat{SR}} = 16.9\%$). Similarly, rather than investing in asset three alone ($\widehat{SR} = 0.459$ and $\hat{\sigma}_{\widehat{SR}} = 13.7\%$), one could invest with portfolio weights $\{w_1 = 60\%, w_2 = 17\%, w_3 = 23\%\}$ realizing $\widehat{SR} = 0.650$ with the same Sharpe ratio standard error ($\hat{\sigma}_{\widehat{SR}} = 13.7\%$).

The portfolio that delivers the highest probabilistic Sharpe ratio has weights $\{w_1 = 61\%, w_2 = 16\%, w_3 = 23\%\}$ realizing $\widehat{SR} = 0.646$ with standard error $\hat{\sigma}_{\widehat{SR}} = 13.6\%$. This portfolio is pointed out in Figure 2.

Conventional portfolio theory would suggest that one should choose the portfolio with the highest Sharpe ratio for a given level of risk. The volatility of expected excess returns of the portfolio (P^*) that delivers the highest probabilistic Sharpe ratio is $\hat{\sigma}(P^*) = 0.96\%$. For this level of uncertainty in excess returns, conventional wisdom would have advocated the portfolio with weights $\{w_1 = 56\%, w_2 = 24\%, w_3 = 20\%\}$ which has $\widehat{SR} = 0.661$ and $\hat{\sigma}_{\widehat{SR}} = 14.5\%$. This portfolio is denoted in Figure 2 with a black circle. This marginally higher Sharpe ratio is significantly less well determined than that of the highest PSR portfolio.

In Figure 3 we plot the expected excess return as a function of the volatility of excess returns. We denote the highest PSR portfolio with a black triangle and the highest Sharpe ratio portfolio for the same level of risk ($\hat{\sigma} = 0.96\%$) with a black square.

Figure 3: Expected excess monthly returns as a function of volatility of excess monthly returns



References

- D. Bailey and M. Lopez de Prado. The Sharpe ratio efficient frontier. 2011. Working Paper.
- S. Christie. Is the Sharpe ratio useful in asset allocation? 2005. MAFC Research Papers No. 31, Applied Finance Centre, Macquarie University.
- A. W. Lo. Statistics of Sharpe ratios. *Financial Analysts Journal*, 58:36–52, 2002.
- H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.
- E. Mertens. Comments on variance of the IID estimator in Lo (2002). 2002. Working paper.
- J. Opdyke. Comparing Sharpe ratios: so where are the p-values? *Journal of Asset Management*, 8(5):308–336, 2007.
- A. D. Roy. Safety first and the holding of assets. *Econometrica*, 20:431–449, 1952.
- W. Sharpe. Mutual fund performance. *Journal of Portfolio Management*, 1:29–34, 1966.

A Central Moments

Let z be a random variable with a given distribution and let $E[f(z)]$ represent the average value of the function $f(z)$ given the distribution. The i^{th} central moment is given by

$$m_i = E[(z - E[z])^i].$$

The average value of z is given by $\mu = E[z]$. One can find simplified expressions for the central moments using the identity $E[E[f(z)]] = f(z)$. We show how this simplifies the first few central moments below

$$\begin{aligned} E[(z - \mu)^2] &= E[z^2] - \mu^2, \\ E[(z - \mu)^3] &= E[z^3] - 3E[z]^2\mu + 2\mu^3, \\ E[(z - \mu)^4] &= E[z^4] - 4E[z^3]\mu + 6E[z^2]\mu^2 - 3\mu^4. \end{aligned}$$

The standard deviation of a distribution is defined by

$$\sigma = E[(z - \mu)^2]^{1/2},$$

the skewness is defined by

$$\gamma_3 = \frac{E[(z - \mu)^3]}{E[(z - \mu)^2]^{3/2}},$$

and the kurtosis is defined by

$$\gamma_4 = \frac{E[(z - \mu)^4]}{E[(z - \mu)^2]^2}.$$