Output Analysis and Run-Length Control

In these notes we describe how the Central Limit Theorem can be used to construct approximate $(1-\alpha)\%$ confidence intervals for the quantity, θ, we are trying to estimate. We also describe methods to estimate the number of samples that are required to achieve a given confidence level and we end with a discussion of the bootstrap method for performing output analysis.

1 Output Analysis

Recall the simulation framework that we use when we want to estimate $\theta := \mathbb{E}[h(\mathbf{X})]$ where $\mathbf{X} \in \mathbb{R}^n$. We first simulate $\mathbf{X}_1, \ldots, \mathbf{X}_n$ IID and then set

$$\widehat{\theta}_n = \frac{h(\mathbf{X}_1) + \ldots + h(\mathbf{X}_n)}{n}$$

The Strong Law of Large Numbers (SLLN) then implies

$$\widehat{\theta}_n \rightarrow \theta \text{ as } n \rightarrow \infty \text{ w.p. } 1.$$

But at this point we don't know how large n should be so that we can have confidence in $\widehat{\theta}_n$ as an estimator of θ. Put another way, for a fixed value of n, what can we say about the quality of $\widehat{\theta}_n$? We will now answer this question and to simplify our notation we will take $Y_i := h(\mathbf{X}_i)$.

1.1 Confidence Intervals

One way to answer this question is to use a *confidence interval*. Suppose then that we want to estimate θ and we have a random vector $\mathbf{Y} = (Y_1, \ldots, Y_n)$ whose distribution depends on θ. Then we seek $L(\mathbf{Y})$ and $U(\mathbf{Y})$ such that

$$\mathbf{P}\left(L(\mathbf{Y}) \leq \theta \leq U(\mathbf{Y})\right) = 1 - \alpha$$

where $0 \leq \alpha \leq 1$ is a pre-specified number. We then say that $[L(\mathbf{Y}), U(\mathbf{Y})]$ is a $100(1-\alpha)\%$ confidence interval for θ. Note that $[L(\mathbf{Y}), U(\mathbf{Y})]$ is a random interval. However, once we replace $\mathbf{Y}$ with a sample vector, $\mathbf{y}$, then $[L(\mathbf{y}), U(\mathbf{y})]$ becomes a real interval. We now discuss the Chebyshev Inequality and the Central Limit Theorem, both of which can be used to construct confidence intervals.

The Chebyshev Inequality

Since the Y_i's are assumed to be IID we know the variance of $\widehat{\theta}_n$ is given by $\mathrm{Var}(\widehat{\theta}_n) = \frac{\sigma^2}{n}$ where $\sigma^2 := \mathrm{Var}(Y)$. Clearly a small value of $\mathrm{Var}(\widehat{\theta}_n)$ implies a more accurate estimate of θ and this is indeed confirmed by *Chebyshev's Inequality* which for any $k > 0$ states that

$$\mathbf{P}\left(|\widehat{\theta}_n - \theta| \geq k\right) \leq \frac{\mathrm{Var}(\widehat{\theta}_n)}{k^2}. \tag{1}$$

We can see from (1) that a smaller value of $\mathrm{Var}(\widehat{\theta}_n)$ therefore improves our confidence in $\widehat{\theta}_n$. We could easily use Chebyshev's Inequality to construct (how?) confidence intervals for θ but it is generally very conservative.

Exercise 1 *Why does Chebyshev's Inequality generally lead to conservative confidence intervals?*

Instead, we will use the Central Limit Theorem to obtain better estimates of $\mathbf{P}\left(|\widehat{\theta}_n - \theta| \geq k\right)$ and as a result, narrower confidence intervals for θ.

The Central Limit Theorem

The Central Limit Theorem is among the most important theorems in probability theory and we state it here for convenience with the symbol "$\stackrel{d}{\rightarrow}$" denoting convergence in distribution.

Theorem 1 (Central Limit Theorem)
Suppose $Y_1, \ldots, Y_n$ are IID and $\mathbb{E}[Y_i^2] < \infty$. Then

$$\frac{\widehat{\theta}_n - \theta}{\sigma/\sqrt{n}} \stackrel{d}{\rightarrow} N(0,1) \quad \text{as} \quad n \rightarrow \infty$$

where $\widehat{\theta}_n = \sum_{i=1}^n Y_i/n$, $\theta := \mathbb{E}[Y_i]$ and $\sigma^2 := \mathrm{Var}(Y_i)$. □

Note that we assume nothing about the distribution of the Y_i's other than that $\mathbb{E}[Y_i^2] < \infty$. If n is sufficiently large in our simulations, then we can use the CLT to construct confidence intervals for $\theta := \mathbb{E}[Y]$. We now describe how to do this.

1.2 An Approximate $100(1-\alpha)\%$ Confidence Interval for θ

Let $z_{1-\alpha/2}$ be the the $(1-\alpha/2)$ percentile point of the $\mathrm{N}(0,1)$ distribution so that

$$\mathbf{P}(-z_{1-\alpha/2} \le Z \le z_{1-\alpha/2}) = 1 - \alpha$$

when $Z \sim \mathrm{N}(0,1)$. Suppose now that we have simulated IID samples, Y_i, for $i = 1, \ldots, n$, and that we want to construct a $100(1-\alpha)\%$ CI for $\theta = \mathbb{E}[Y]$. That is, we want $L(\mathbf{Y})$ and $U(\mathbf{Y})$ such that

$$\mathbf{P}\left(L(\mathbf{Y}) \le \theta \le U(\mathbf{Y})\right) = 1 - \alpha.$$

The CLT implies $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)/\sigma$ is approximately $\mathrm{N}(0,1)$ for large n so we have

$$\begin{aligned}
\mathbf{P}\left(-z_{1-\alpha/2} \;\le\; \frac{\sqrt{n}(\widehat{\theta}_n - \theta)}{\sigma} \;\le\; z_{1-\alpha/2}\right) &\approx 1-\alpha \\
\Rightarrow\; \mathbf{P}\left(-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \;\le\; \widehat{\theta}_n - \theta \;\le\; z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &\approx 1-\alpha \\
\Rightarrow\; \mathbf{P}\left(\widehat{\theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \;\le\; \theta \;\le\; \widehat{\theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &\approx 1-\alpha.
\end{aligned}$$

Our approximate $100(1-\alpha)\%$ CI for θ is therefore given by

$$[L(\mathbf{Y}),\ U(\mathbf{Y})] \;=\; \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\;\; \widehat{\theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]. \tag{2}$$

Recall that $\widehat{\theta}_n = (Y_1 + \ldots + Y_n)/n$, so L and U are indeed functions of $\mathbf{Y}$. There is still a problem, however, as we do not usually know σ^2. We resolve this issue by estimating σ^2 with

$$\widehat{\sigma}_n^2 = \frac{\sum_{i=1}^n (Y_i - \widehat{\theta}_n)^2}{n-1}.$$

It is easy to show that $\widehat{\sigma}_n^2$ is an unbiased estimator of σ^2 and that $\widehat{\sigma}_n^2 \rightarrow \sigma^2$ w.p. 1 as $n \rightarrow \infty$. So now we replace σ with $\widehat{\sigma}_n$ in (2) to obtain

$$[L(\mathbf{Y}),\ U(\mathbf{Y})] = \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}},\;\; \widehat{\theta}_n + z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\right] \tag{3}$$

as our *approximate* $100(1-\alpha)\%$ CI for θ when n is large.

Remark 1 *Note that when we obtain sample values of* $\mathbf{y} = (y_1, \ldots, y_n)$*, then* $[L(\mathbf{y}),\ U(\mathbf{y})]$ *becomes a real interval. Then we can no longer say (why not?) that*

$$\mathbf{P}\left(\theta \in [L(\mathbf{y}),\ U(\mathbf{y})]\right) = 1 - \alpha.$$

Instead, we say that we are $100(1-\alpha)\%$ confident *that* $[L(\mathbf{y}),\ U(\mathbf{y})]$ *contains* θ*. Furthermore, the smaller the value of* $U(\mathbf{y}) - L(\mathbf{y})$*, the more confidence we will have in our estimate of* θ*.*

Example 1 (Pricing a European Call Option)
Suppose we want to estimate the price, C_0, of a call option on a stock whose price process, S_t, is a $GBM(\mu, \sigma)$. The relevant parameters are $r = .05$, $T = 0.5$ years, $S_0 = \$100$, $\sigma = 0.2$ and strike $K = \$110$. Then we know that

$$C_0 \ = \ \mathbb{E}_0^Q[e^{-rT}\max(S_T - K, 0)]$$

where we can assume that $S_t \sim GBM(r, \sigma)$ under the risk-neutral probability measure, Q. That is, we assume $S_T = S_0 \exp\left((r - \sigma^2/2)T + \sigma Z\right)$ where $Z \sim \mathsf{N}(0, T)$. Though we can of course compute C_0 exactly, we can also estimate C_0 using Monte Carlo with (3) yielding an approximate $100(1-\alpha)\%$ CI for C_0 with $Y_i := e^{-rT}\max(S_T^{(i)} - K, 0)$ denoting the i^{th} discounted sample payoff of the option. Based on $n = 100$k samples, we obtain $[15.16,\ 15.32]$ as our approximate 95% CI for C_0. ∎

Properties of the Confidence Interval

The *width* of the confidence interval is given by

$$U - L = \frac{2\widehat{\sigma}_n z_{1-\alpha/2}}{\sqrt{n}}$$

and so the *half-width* then is $(U - L)/2$. The width clearly depends on α, $\widehat{\sigma}_n$ and n. However, $\widehat{\sigma}_n \rightarrow \sigma$ almost surely as $n \rightarrow \infty$, and σ is a constant. Therefore, for a fixed α, we need to increase n if we are to decrease the width of the confidence interval. Indeed, since $U - L \propto \frac{1}{\sqrt{n}}$, we can see for example that we would need to increase n by a factor of four in order to decrease the width of the confidence interval by only a factor of two.

---

2 Run-Length Control

Up to this point we have selected n in advance and then computed the approximate CI. The width of the CI is then a measure of the error in our estimator. Now we will do the reverse by first choosing some error criterion that we want our estimator to satisfy, and then choosing n so that this criterion is satisfied.

There are two types of error that we will consider:

1. *Absolute error*, which is given by $E_a := |\widehat{\theta}_n - \theta|$ and
2. *Relative error*, which is given by $E_r := \left|\frac{\widehat{\theta}_n - \theta}{\theta}\right|$.

Now we know that $\widehat{\theta}_n \rightarrow \theta$ w.p. 1 as $n \rightarrow \infty$ so that E_a and E_r both $\rightarrow 0$ as $n \rightarrow \infty$. (If $\theta = 0$ then E_r is not defined.) However, in practice $n \neq \infty$ and so the errors will be non-zero. We specify the following error criterion:

Error Criterion: Given $0 \leq \alpha \leq 1$ and $\epsilon \geq 0$, we want $\mathbf{P}(E \leq \epsilon) = 1 - \alpha$. E is the error type we have specified, i.e., relative or absolute.

The goal then is to choose n so that the error criterion is approximately satisfied and this is easily done. Suppose, for example, that we want to control absolute error, E_a. Then, as we saw earlier,

$$\mathbf{P}\left(\widehat{\theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \ \leq \ \theta \ \leq \ \widehat{\theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \approx 1 - \alpha.$$

This then implies $\mathbf{P}\left(|\widehat{\theta}_n - \theta| \leq z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \approx 1-\alpha$, so in terms of E_a we have

$$\mathbf{P}\left(E_a \leq z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \approx 1-\alpha.$$

If we then want $\mathbf{P}\left(E_a \leq \epsilon\right) \approx 1-\alpha$, it clearly suffices to choose n such that

$$n = \frac{\sigma^2 z_{1-\alpha/2}^2}{\epsilon^2}.$$

If we are working with relative error, then a similar argument implies that $\mathbf{P}\left(E_r \leq \epsilon\right) \approx 1-\alpha$ if

$$n = \frac{\sigma^2 z_{1-\alpha/2}^2}{\theta^2 \epsilon^2}.$$

There are still some problems, however:

1. When we are controlling E_r, we need to know σ and θ in advance.
2. When we are controlling E_a, we need to know σ in advance.

Of course we do not usually know σ or θ in advance. In fact, θ is what we are trying to estimate! There are two methods we can use to overcome this problem: the *two-stage method* and the *sequential method*, both of which we will now describe.

2.1 The Two-Stage Procedure

Suppose we want to satisfy the condition $\mathbf{P}(E_a \leq \epsilon) = 1-\alpha$ so that we are trying to control the absolute error. Then we saw earlier that we would like to set

$$n = \frac{\sigma^2 z_{1-\alpha/2}^2}{\epsilon^2}.$$

Unfortunately, we don't know σ^2 but we can solve this problem by first doing a *pilot* simulation to estimate it. The idea is to do a small number, p, of initial runs to estimate σ^2. We then use our estimate, $\widehat{\sigma}^2$, to compute an estimate, $\widehat{n}$, of n. Finally, we repeat the simulation, but now we use $\widehat{n}$ runs. We have the following algorithm.

Two-Stage Monte Carlo Simulation for Estimating $\mathbb{E}[h(\mathbf{X})]$

/∗ Do pilot simulation first ∗/
for $i = 1$ **to** p
    generate $\mathbf{X}^i$
end for
set $\widehat{\theta} = \sum h(\mathbf{X}^i)/p$
set $\widehat{\sigma}^2 = \sum(h(\mathbf{X}^i) - \widehat{\theta})^2/(p-1)$
set $n = \frac{\widehat{\sigma}^2 z_{1-\alpha/2}^2}{\epsilon^2}$

/∗ Now do main simulation ∗/
for $i = 1$ **to** n
    generate $\mathbf{X}^i$
end for
set $\widehat{\theta}_n = \sum h(\mathbf{X}^i)/n$
set $\widehat{\sigma}_n^2 = \sum(h(\mathbf{X}^i) - \widehat{\theta}_n)^2/(n-1)$
set $100(1-\alpha)$ % CI $= \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\ ,\ \widehat{\theta}_n + z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\right]$

For this method to work, it is important that $\widehat{\theta}$ and $\widehat{\sigma}^2$ be sufficiently good estimates of θ and σ^2. Therefore, it is important to make p sufficiently large. In practice, we usually take $p \geq 50$. We can use an analogous two-stage procedure if we want to control the *relative* error and have $\mathbf{P}(E_r \leq \epsilon) = 1 - \alpha$.

2.2 The Sequential Procedure

Suppose again that we wish to satisfy the condition $\mathbf{P}(E_a \leq \epsilon) = 1 - \alpha$. Then we saw earlier that we would like to set

$$n = \frac{\sigma^2 z_{1-\alpha/2}^2}{\epsilon^2}.$$

In contrast to the pilot procedure, we do not precompute n during the sequential procedure. Instead, we continue to generate samples until

$$\frac{\widehat{\sigma}_n z_{1-\alpha/2}}{\sqrt{n}} \leq \epsilon$$

where $\widehat{\sigma}_n$ is again the estimate of σ based upon the first n samples. It is important that n be sufficiently large so that $\widehat{\theta}_n$ and $\widehat{\sigma}_n^2$, are sufficiently good estimates of θ and σ^2, respectively. As a result, we typically insist that $n \geq 50$ before we stop. Approximate confidence intervals are then computed as usual.

Question: Have we allowed any biases to creep in here?

We have the following algorithm:

Sequential Monte Carlo Simulation for Estimating $\mathbb{E}[h(\mathbf{X})]$

> **set** check $= 0,\ n = 1$
> **while** (check $= 0$)
>
>   **generate** $\mathbf{X}^n$
>   **set** $\widehat{\theta}_n = \sum h(\mathbf{X}^i)/n$
>   **set** $\widehat{\sigma}_n^2 = \sum (h(\mathbf{X}^i) - \widehat{\theta}_n)^2/(n-1)$
>   **if** $(n \geq p)$ and $\left(\frac{\widehat{\sigma}_n z_{1-\alpha/2}}{\sqrt{n}} \leq \epsilon\right)$
>   check $= 1$
>   **else**
> $n = n + 1$
>   **end if**
> **end while**
> **set** $100(1-\alpha)$ % CI $= \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\ ,\ \widehat{\theta}_n + z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\right]$

In practice we do **not** need to store every value, $h(\mathbf{X}^i)$ for $i = 1, \ldots, n$, in order to update $\widehat{\theta}_n$ and $\widehat{\sigma}_n$. Indeed, we can update $\widehat{\theta}_n$ and $\widehat{\sigma}_n$ efficiently by observing that

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} + \frac{h(\mathbf{X}^n) - \widehat{\theta}_{n-1}}{n} \quad \text{and}$$

$$\widehat{\sigma}_n^2 = \left(\frac{n-2}{n-1}\right)\widehat{\sigma}_{n-1}^2 + n\left(\widehat{\theta}_n - \widehat{\theta}_{n-1}\right)^2.$$

If we want to control the *relative* error and have $\mathbf{P}(E_r \leq \epsilon) = 1 - \alpha$, then we would simulate samples until

$$\frac{\widehat{\sigma}_n(z_{1-\alpha/2})}{\widehat{\theta}_n\sqrt{n}} \leq \epsilon.$$

3 Output Analysis Using the Bootstrap

We can view[1] our output analysis problem as one of estimating

$$\mathsf{MSE}(F) := \mathbb{E}_F\left[(g(Y_1,\ldots,Y_n)-\theta(F))^2\right] \tag{4}$$

where $\theta(F)=\mathbb{E}_F[X]$, $g(Y_1,\ldots,Y_n):=\bar{Y}$ and F denotes the CDF of Y. In that case, we saw in Section 1 how we could use the CLT to construct approximate confidence intervals for θ. While this is certainly the most common context in which we encounter (4), other situations arise where the CLT cannot be easily used to obtain a confidence interval for $\theta(F)$. For example, if $\theta(F)=\mathrm{Var}(Y)$ or $\theta(F)=\mathbb{E}\left[Y \mid Y\geq\alpha\right]$, then an alternative method of constructing a confidence interval for θ will be required. The bootstrap method provides such an alternative and in order to describe the method we will assume our problem is to estimate $\mathsf{MSE}(F)$ as in (4).

To begin with, recall that the empirical distribution, F_e, is defined to be the CDF of the distribution that places a weight of $1/n$ on each of the simulated values $Y_1,\ldots,Y_n$. The empirical CDF therefore satisfies

$$F_e(y)=\frac{\sum_{i=1}^n 1_{\{Y_i\leq y\}}}{n}$$

and for large n it can be shown (and should be intuitively clear) that F_e should be[2] a good approximation to F. Therefore, as long as θ is sufficiently well-behaved, i.e. a "continuous" function of F, then for sufficiently large n we should have

$$\mathsf{MSE}(F)\approx\mathsf{MSE}(F_e)=\mathbb{E}_{F_e}\left[(g(Y_1,\ldots,Y_n)-\theta(F_e))^2\right]. \tag{5}$$

The quantity $\mathsf{MSE}(F_e)$ is known as the *bootstrap approximation* to $\mathsf{MSE}(F)$ and is easy to estimate via simulation as we shall see below. But first, however, we will consider an example where $\mathsf{MSE}(F_e)$ can be computed exactly. Indeed the bootstrap is not required in this case but it is nonetheless instructive to see the calculations written out explicitly.

Example 2 (Applying the Bootstrap to the Sample Mean)
Suppose we wish to estimate $\theta(F)=\mathbb{E}_F[Y]$ via the estimator $\widehat{\theta}=g(Y_1,\ldots,Y_n):=\bar{Y}$. As noted above, the bootstrap is not necessary in this case as we can apply the CLT directly as in Section 1 to obtain confidence intervals for $\widehat{\theta}$ or equivalently, we can estimate the mean-squared error $\mathbb{E}\left[\left(\bar{Y}-\theta\right)^2\right]=\sigma^2/n$ with $\widehat{\sigma}_n^2/n=\sum_{i=1}^n(y_i-\bar{y})^2/(n(n-1))$.

Letting $\bar{y}$ denote the mean of the observed, i.e. simulated, data-points $y_1,\ldots,y_n$, we obtain that the bootstrap estimator is given by

$$\begin{aligned}
\mathsf{MSE}(F_e) &= \mathbb{E}_{F_e}\left[\left(\frac{\sum_{i=1}^n Y_i}{n}-\bar{y}\right)^2\right] \\
&= \mathrm{Var}_{F_e}\left(\frac{\sum_{i=1}^n Y_i}{n}\right) && (6)\\
&= \frac{\mathrm{Var}_{F_e}(Y)}{n} && (7)\\
&= \frac{\sum_{i=1}^n (y_i-\bar{y})^2}{n^2}
\end{aligned}$$

where (6) follows since $\mathbb{E}_{F_e}[Y]=\bar{y}$, and (7) follows since the Y_i's are IID F_e. We therefore see that the bootstrap approximation to the MSE is almost identical to our usual estimator, $\widehat{\sigma}_n^2/n$. ∎

[1]We follow Sheldon M. Ross's *Simulation* in our development of the bootstrap here.
[2]Indeed it can be shown that $F_e(y)$ converges to $F(y)$ uniformly in y w.p. 1 as $n\to\infty$.

In contrast to Example 2, we cannot usually compute MSE(F_e) explicitly, but as it's an expectation we can easily use Monte-Carlo to estimate it. In this case we need to simulate from F_e which is easy to do and so we obtain the following bootstrap algorithm for estimating MSE(F).

Bootstrap Simulation Algorithm for Estimating MSE(F)

```
for i = 1 to B
    generate Y_1, ..., Y_n IID from F_e
    set θ̂_i^b = g(Y_1, ..., Y_n)
    set Z_i^b = [θ̂_i^b − θ(F_e)]^2
end for
set MSE(F)^ = Σ_{b=1}^{B} Z^(b)/B
```

for $i = 1$ **to** B
  generate $Y_1, \ldots, Y_n$ IID from F_e
  set $\widehat{\theta}_i^b = g(Y_1, \ldots, Y_n)$
  set $Z_i^b = \left[\widehat{\theta}_i^b - \theta(F_e)\right]^2$
end for
set $\widehat{\text{MSE}(F)} = \sum_{b=1}^{B} Z^{(b)}/B$

The Z_i^b's (or equivalently the $\widehat{\theta}_i^b$'s) are the bootstrap samples and a value of $B = 100$ is often sufficient to obtain a sufficiently accurate estimate. In the next example we apply the bootstrap approach in a *historical* simulation context where we have real data observations as opposed to simulated data. (The disadvantage with historical simulation is that we typically have no control over n.)

Example 3 (Estimating the Minimum Variance Portfolio)
Suppose we wish to invest a fixed sum of money in two financial assets, X and Z say, that yield random returns of R_x and R_z, respectively. We invest a fraction θ of our wealth in X, and the remaining $1-\theta$ in Z. The goal is to choose θ to minimize the total variance, Var($\theta R_x + (1-\theta)R_z$), of our investment return. It is easy to see that the minimizing θ is given by

$$\theta = \frac{\sigma_z^2 - \sigma_{xz}}{\sigma_x^2 + \sigma_z^2 - 2\sigma_{xz}} \tag{8}$$

where $\sigma_x^2 = \mathrm{Var}(R_x)$, $\sigma_z^2 = \mathrm{Var}(R_z)$ and $\sigma_{xz} = \mathrm{Cov}(R_x, R_z)$. In practice, we do not know these quantities and therefore have to estimate them from historical data. We therefore obtain

$$\widehat{\theta} = \frac{\widehat{\sigma}_z^2 - \widehat{\sigma}_{xz}}{\widehat{\sigma}_x^2 + \widehat{\sigma}_z^2 - 2\widehat{\sigma}_{xz}}. \tag{9}$$

as our estimator of the minimum variance portfolio with $\widehat{\sigma}_x^2$, $\widehat{\sigma}_z^2$ and $\widehat{\sigma}_{xz}$ estimated from historical return data $Y_1, \ldots, Y_n$ with $Y_i := \left(R_x^{(i)}, R_z^{(i)}\right)$ the joint return in period i.

We would like to know how good an estimator $\widehat{\theta}$ is. More specifically, what is the (mean-squared) error when we use $\widehat{\theta}$? We can answer this question using the bootstrap with $\theta(F) := \theta$ and $g(Y_1, \ldots, Y_n) = \widehat{\theta}$ the estimator given by (9). ∎

Exercise 2 *Provide pseudo-code for estimating MSE($\widehat{\theta}$) := MSE(F), in Example 3.*

Exercise 3 *Consider the problem of estimating $\theta(F) = \mathbb{E}[Y \mid Y \geq \beta]$ for some fixed constant, β. Explain how you would use the bootstrap to estimate MSE(F) in this case given n Monte-Carlo samples $Y_1, \ldots, Y_n$.*

3.1 Constructing Bootstrap Confidence Intervals

The bootstrap method is also widely used to construct confidence intervals and here we will consider the so-called *basic bootstrap interval*. Consider our bootstrap samples $\widehat{\theta}_1^b, \ldots, \widehat{\theta}_B^b$ and suppose we want a $1-\alpha$ confidence interval for $\theta = \theta(F)$. Let q_l and q_u be the $\alpha/2$ lower- and upper-sample quantiles, respectively, of the bootstrap samples. Then the fraction of bootstrap samples satisfying

$$q_l \leq \widehat{\theta}^b \leq q_u \tag{10}$$

is $1-\alpha$. But (10) is equivalent to

$$\widehat{\theta} - q_u \leq \widehat{\theta} - \widehat{\theta}^b \leq \widehat{\theta} - q_l \tag{11}$$

where $\widehat{\theta} = g(y_1, \ldots, y_n)$ is our estimate of θ computed using the original data-set. This implies $\widehat{\theta} - q_u$ and $\widehat{\theta} - q_l$ are the lower and upper quantiles for $\widehat{\theta} - \widehat{\theta}^b$. The basic bootstrap assumes they are also the quantiles for $\theta - \widehat{\theta}$. This makes sense intuitively – and can be justified mathematically as $n \to \infty$ and if θ is a "continuous" function of F. It therefore follows that

$$\widehat{\theta} - q_u \leq \theta - \widehat{\theta} \leq \widehat{\theta} - q_l \tag{12}$$

will occur in approximately in a fraction $1-\alpha$ of samples. Adding $\widehat{\theta}$ across (12) yields an approximate $(1-\alpha)\%$ CI for θ of

$$(2\widehat{\theta} - q_u,\ 2\widehat{\theta} - q_l).$$