Portfolios that Contain Risky Assets 11: Independent, Identically-Distributed Models

C. David Levermore

University of Maryland, College Park, MD

Math 420: *Mathematical Modeling*
March 21, 2018 version

Portfolios that Contain Risky Assets
Part II: Stochastic Models

Independent, Identically-Distributed Models

1. Independent, Identically-Distributed Models for Assets
2. Return Probability Densities
3. Growth Rate Probability Densities
4. Normal Growth Rate Model
5. Independent, Identically-Distributed Models for Markets
6. Independent, Identically-Distributed Models for Portfolios

Independent, Identically-Distributed Models for Assets

Investors have long followed the old adage "don't put all your eggs in one basket" by holding diversified portfolios. However, before MPT the value of diversification had not been quantified. Key aspects of MPT are:

1. it uses the return mean as a proxy for reward;
2. it uses volatility as a proxy for risk;
3. it analyzes Markowitz portfolios;
4. it shows diversification reduces volatility through covariances;
5. it identifies the efficient frontier as the place to be.

Independent, Identically-Distributed Models for Assets

The orignial form of MPT did not give guidance to investors about where to be on the efficient frontier. This question was addressed in the 1960's, most notably by William Sharpe, who shared the 1990 Nobel Prize in Economics with Harry Markowitz. We will not present that work here. Rather, we will build stochastic models that can be used in conjunction with MPT to address this question. *By doing so, we will learn that maximizing the return mean is not the best strategy for maximizing reward.*

We begin by building models of one risky asset with a share price history $\{s(d)\}_{d=0}^{D}$. Let $\{r(d)\}_{d=1}^{D}$ be the associated return history. Because each $s(d)$ is positive, each $r(d)$ lies in the interval $(-1, \infty)$.

Independent, Identically-Distributed Models for Assets

An *independent, identically-distributed (IID)* model for this history simply independently draws D random numbers $\{R(d)\}_{d=1}^{D}$ from $(-1,\infty)$ in accord with a fixed probability density $q(R)$ over $(-1,\infty)$. This means that $q(R)$ is a nonnegative integrable function such that

$$\int_{-1}^{\infty} q(R)\,\mathrm{d}R = 1\,,$$

and that the probability that each $R(d)$ takes a value inside any interval $[R_1,R_2] \subset (-1,\infty)$ is given by

$$\mathrm{Pr}\{R(d) \in [R_1,R_2]\} = \int_{R_1}^{R_2} q(R)\,\mathrm{d}R\,.$$

Here capital letters $R(d)$ denote random numbers drawn from $(-1,\infty)$ in accord with the probability density $q(R)$ rather than real return data.

Independent, Identically-Distributed Models for Assets

Because the random numbers $\{R(d)\}_{d=1}^{D}$ are drawn from $(-1, \infty)$ in accord with the probability density $q(R)$ *independent* of each other, there is no correlation of $R(d)$ with $R(d')$ when $d \neq d'$. In particular, if we plot the points $\{(R(d), R(d+c))\}_{d=1}^{D-c}$ in the rr'-plane for any $c > 0$ they will be distributed in accord with the probability density $q(r)q(r')$. *Therefore if the return history $\{r(d)\}_{d=1}^{D}$ is mimiced by such a model then the points $\{(r(d), r(d+c))\}_{d=1}^{D-c}$ plotted in the rr'-plane should appear to be distributed in a way consistant with the probability density $q(r)q(r')$.*

Such plots are called *scatter plots.* In general a scatter plot will not show independence when c is small. This is because the behavior of an asset on any given trading day generally correlates with its behavior on the previous trading day. However, if a scatter plot shows independence for some c that is small compared to D the an IID model might still be good. Such a time c is called the *correlation time.*

Independent, Identically-Distributed Models for Assets

Because the random numbers $\{R(d)\}_{d=1}^{D}$ are each drawn from $(-1,\infty)$ in accord with the *same* probability density $q(R)$, if we plot the points $\{(d,R(d))\}_{d=1}^{D}$ in the dr-plane they will usually be distributed in a way that looks uniform in d. *Therefore if the return history $\{r(d)\}_{d=1}^{D}$ is mimiced by such a model then the points $\{(d,r(d))\}_{d=1}^{D}$ plotted in the dr-plane should appear to be distributed in a way that is unifrom in d.*

Exercise. Plot $\{(r(d),r(d+1))\}_{d=1}^{D-1}$ and $\{(d,r(d))\}_{d=1}^{D}$ for each of the following assets and explain which might be good candidates to be mimiced by an IID model.

(a) Google, Microsoft, Exxon-Mobil, UPS, GE, and Ford stock in 2017;

(b) Google, Microsoft, Exxon-Mobil, UPS, GE, and Ford stock in 2008;

(c) S&P 500 and Russell 1000 and 2000 index funds in 2017;

(d) S&P 500 and Russell 1000 and 2000 index funds in 2008.

Independent, Identically-Distributed Models for Assets

Remark. We have adopted IID models because they are simple. It is not hard to develop more complicated stochastic models. For example, we could use a different probability density for each day of the week rather than treating all trading days the same way. Because there are usually five trading days per week, Monday through Friday, such a model would require calibrating five times as many means and covariances with one fifth as much data. There would then be greater uncertainty associated with the calibration. Moreover, we then have to figure out how to treat weeks that have less than five trading days due to holidays. Perhaps just the first and last trading days of each week should get their own probability density, no matter on which day of the week they fall. *Before increasing the complexity of a model, you should investigate whether the costs of doing so outweigh the benefits.* Specifically, you should investigate whether or not there is benefit in treating any one trading day of the week differently than the others before building a more complicated models.

Independent, Identically-Distributed Models for Assets

Remark. *IID models are the simplest models that are consistent with the way any portfolio theory is used.* Specifically, to use any portfolio theory we must first calibrate a model from historical data. This model is then used to suggest how a set of ideal portfolios might behave in the future. Based on these suggestions we select the ideal portfolio that optimizes some objective. *This strategy assumes that in the future the market will behave statistically as it did in the past.*

This assumption requires the market statistics to be stable relative to its dynamics. But this requires future states to decorrelate from past states. Markov models are characterized by the assumption that possible future states are independent of past states, which maximizes this decorrelation. IID models are the simplest Markov models. All the models discussed in the previous remark are also Markov models. We will use only IID models.

Return Probability Densities

Once we have decided to use an IID model for a particular asset, you might think the next goal is to pick an appropriate probability density $q(R)$. However, that approach is neither practical nor necessary.

Rather, the goal is to identify appropriate statistical information about $q(R)$ that sheds light on the market. Ideally this information should be insensitive to details of $q(R)$ within a large class of probability densities. Statisticians call such an approach *nonparametric.*

Return Probability Densities

The expected value of any function $\psi(R)$ is given by

$$\mathrm{Ex}(\psi(R)) = \int_{-1}^{\infty} \psi(R)\, q(R)\, \mathrm{d}R\,,$$

provided $|\psi(R)|\, q(R)$ is integrable. Because we have been computing return means and variances, we will assume that $q(R)$ satisfies

$$\int_{-1}^{\infty} R^2 q(R)\, \mathrm{d}R < \infty\,.$$

The mean μ and variance ξ of R are then

$$\mu = \mathrm{Ex}(R) = \int_{-1}^{\infty} R\, q(R)\, \mathrm{d}R\,,$$

$$\xi = \mathrm{Var}(R) = \mathrm{Ex}\Big((R-\mu)^2\Big) = \int_{-1}^{\infty} (R-\mu)^2\, q(R)\, \mathrm{d}R\,.$$

Return Probability Densities

Because μ and ξ are not known, they must be estimated from the data.

Given D samples $\{R(d)\}_{d=1}^{D}$ that are drawn from the density $q(R)$, we can construct an *estimator* $\hat{\mu}$ of μ by

$$\hat{\mu} = \sum_{d=1}^{D} w(d)\, R(d)\,.$$

This so-called *sample mean* is an *unbiased estimator* of μ because

$$\mathrm{Ex}(\hat{\mu}) = \sum_{d=1}^{D} w(d)\, \mathrm{Ex}\big(R(d)\big) = \sum_{d=1}^{D} w(d)\, \mu = \mu\,.$$

Return Probability Densities

We can estimate how close $\hat{\mu}$ is to μ by computing its variance as

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}) &= \mathrm{Ex}\Big((\hat{\mu} - \mu)^2\Big) \\
&= \mathrm{Ex}\left(\sum_{d=1}^{D} \sum_{d'=1}^{D} w(d)\, w(d')\, (R(d) - \mu)\, (R(d') - \mu)\right) \\
&= \sum_{d=1}^{D} \sum_{d'=1}^{D} w(d)\, w(d')\, \mathrm{Ex}\big((R(d) - \mu)\, (R(d') - \mu)\big) \\
&= \sum_{d=1}^{D} w(d)^2\, \mathrm{Ex}\big((R(d) - \mu)^2\big) \\
&= \sum_{d=1}^{D} w(d)^2\, \mathrm{Var}(R) = \sum_{d=1}^{D} w(d)^2\, \xi\,.
\end{aligned}$$

Return Probability Densities

Remark. The off-diagonal terms in the foregoing double sum vanish because

$$\mathrm{Ex}\big((R(d)-\mu)\,(R(d')-\mu)\big) = 0 \quad \text{when } d \neq d' \,.$$

We now define $\bar{w}$ by

$$\bar{w} = \sum_{d=1}^{D} w(d)^2 \,.$$

Then the result of the calculation on the last slide can be expressed as

$$\mathrm{Var}(\hat{\mu}) = \bar{w}\,\xi \,.$$

This fact will play important roles in the next two slides.

Return Probability Densities

Remark. The fact $\mathrm{Var}(\hat{\mu}) = \bar{w}\,\xi$ implies that $\hat{\mu}$ converges to μ like $\sqrt{\bar{w}}$ as $D \to \infty$. The Cauchy inequality implies that

$$1 = \left(\sum_{d=1}^{D} w(d) \right)^2 \leq \left(\sum_{d=1}^{D} 1^2 \right) \left(\sum_{d=1}^{D} w(d)^2 \right) = D\,\bar{w}\,.$$

Therefore $1/D \leq \bar{w}$ for any choice of weights. Because $\bar{w} = 1/D$ for uniform weights, we see that the rate of convergence of $\hat{\mu}$ to μ is fastest for uniform weights, when it is $1/\sqrt{D}$ as $D \to \infty$.

We can construct an *unbiased estimator* of ξ that is proportional to the so-called *sample variance* as

$$\hat{\xi} = \frac{1}{1-\bar{w}} \sum_{d=1}^{D} w(d)\big(R(d) - \hat{\mu}\big)^2\,.$$

Return Probability Densities

Indeed, by using the fact that $\mathrm{Var}(\hat{\mu}) = \bar{w}\,\xi$ we confirm that

$$\begin{aligned}
\mathrm{Ex}(\hat{\xi}) &= \frac{1}{1-\bar{w}}\,\mathrm{Ex}\left(\sum_{d=1}^{D} w(d)\big(R(d)-\mu\big)^2 - (\hat{\mu}-\mu)^2\right) \\
&= \sum_{d=1}^{D} \frac{w(d)}{1-\bar{w}}\,\mathrm{Ex}\Big(\big(R(d)-\mu\big)^2\Big) - \frac{\mathrm{Ex}\big((\hat{\mu}-\mu)^2\big)}{1-\bar{w}} \\
&= \sum_{d=1}^{D} \frac{w(d)}{1-\bar{w}}\,\mathrm{Var}(R) - \frac{\mathrm{Var}(\hat{\mu})}{1-\bar{w}} \\
&= \frac{\xi}{1-\bar{w}} - \frac{\bar{w}\,\xi}{1-\bar{w}} = \xi\,.
\end{aligned}$$

Growth Rate Probability Densities

Given D samples $\{R(d)\}_{d=1}^{D}$ that are drawn from the return probability density $q(R)$, the associated simulated share prices satisfy

$$S(d) = (1 + R(d))\, S(d-1)\,, \quad \text{for } d = 1, \cdots, D\,.$$

If we set $S(0) = s(0)$ then you can easily see that

$$S(d) = \prod_{d'=1}^{d} (1 + R(d'))\, s(0)\,.$$

Growth Rate Probability Densities

The *growth rate* $X(d)$ is related to the return $R(d)$ by

$$e^{X(d)} = 1 + R(d) \,.$$

In other words, $X(d)$ is the growth rate that yeilds a return $R(d)$ on trading day d. The formula for $S(d)$ then takes the form

$$S(d) = \exp\left(\sum_{d'=1}^{d} X(d') \right) s(0) \,.$$

Growth Rate Probability Densities

If the samples $\{R(d)\}_{d=1}^{D}$ are drawn from a density $q(R)$ over $(-1,\infty)$ then the $\{X(d)\}_{d=1}^{D}$ are drawn from a density $p(X)$ over $(-\infty,\infty)$ where

$$p(X)\,\mathrm{d}X = q(R)\,\mathrm{d}R\,,$$

with X and R related by

$$X = \log(1+R)\,, \qquad R = e^{X} - 1\,.$$

More explicitly, the densities $p(X)$ and $q(R)$ are related by

$$p(X) = q\Big(e^{X} - 1\Big)\,e^{X}\,, \qquad q(R) = \frac{p(\log(1+R))}{1+R}\,.$$

Growth Rate Probability Densities

Because our models will involve means and variances, we will require that

$$\int_{-\infty}^{\infty} X^2 p(X)\,\mathrm{d}X = \int_{-1}^{\infty} \log(1+R)^2\, q(R)\,\mathrm{d}R < \infty\,,$$
$$\int_{-\infty}^{\infty} \left(e^X - 1\right)^2 p(X)\,\mathrm{d}X = \int_{-1}^{\infty} R^2 q(R)\,\mathrm{d}R < \infty\,.$$

Then the mean γ and variance θ of X are

$$\gamma = \mathrm{Ex}(X) = \int_{-\infty}^{\infty} X\, p(X)\,\mathrm{d}X\,,$$
$$\theta = \mathrm{Var}(X) = \mathrm{Ex}\Big((X-\gamma)^2\Big) = \int_{-\infty}^{\infty} (X-\gamma)^2\, p(X)\,\mathrm{d}X\,.$$

Growth Rate Probability Densities

The big advantage of working with $p(X)$ rather than $q(R)$ is the fact that

$$\log\left(\frac{S(d)}{s(0)}\right) = \sum_{d'=1}^{d} X(d')\,.$$

In other words, $\log(S(d)/s(0))$ *is a sum of an IID process.* It is easy to compute the mean and variance of this quantity in terms of those of X.

For the mean of $\log(S(d)/s(0))$ we find that

$$\mathrm{Ex}\left(\log\left(\frac{S(d)}{s(0)}\right)\right) = \sum_{d'=1}^{d} \mathrm{Ex}\big(X(d')\big) = d\,\gamma\,,$$

Growth Rate Probability Densities

For the variance of $\log(S(d)/s(0))$ we find that

$$\begin{aligned}
\mathrm{Var}\left(\log\left(\frac{S(d)}{s(0)}\right)\right) &= \mathrm{Ex}\left(\left(\sum_{d'=1}^{d} X(d') - d\,\gamma\right)^2\right) \\
&= \mathrm{Ex}\left(\left(\sum_{d'=1}^{d} \big(X(d') - \gamma\big)\right)^2\right) \\
&= \mathrm{Ex}\left(\sum_{d'=1}^{d}\sum_{d''=1}^{d} \big(X(d') - \gamma\big)\big(X(d'') - \gamma\big)\right) \\
&= \sum_{d'=1}^{d} \mathrm{Ex}\Big(\big(X(d') - \gamma\big)^2\Big) = d\,\theta\,.
\end{aligned}$$

Growth Rate Probability Densities

Remark. The off-diagonal terms in the foregoing double sum vanish because

$$\mathrm{Ex}\Big(\big(X(d') - \gamma\big)\ \big(X(d'') - \gamma\big)\Big) = 0 \qquad \text{when } d'' \neq d' \,.$$

Hence, the growth mean and variance of the IID model asset at day d is

$$\mathrm{Ex}\left(\log\left(\frac{S(d)}{s(0)}\right)\right) = \gamma\, d\,, \qquad \mathrm{Var}\left(\log\left(\frac{S(d)}{s(0)}\right)\right) = \theta\, d\,.$$

Growth Rate Probability Densities

Remark. *The IID model suggests that the growth rate mean γ is a good proxy for the reward of an asset and that $\sqrt{\theta}$ is a good proxy for its risk. However, these are not the proxies chosen by MPT when it is applied to a portfolio consisting of one risky asset.*

The proxies γ and $\sqrt{\theta}$ can be approximated by $\hat{\gamma}$ and $\sqrt{\hat{\theta}}$ where $\hat{\gamma}$ and $\hat{\theta}$ are the unbiased estimators of γ and θ given by

$$\hat{\gamma} = \sum_{d=1}^{D} w(d)\, X(d)\,, \qquad \hat{\theta} = \sum_{d=1}^{D} \frac{w(d)}{1-\bar{w}} \left(X(d) - \hat{\gamma}\right)^2 .$$

Normal Growth Rate Model

We can illustrate what is going on with the simple IID model where $p(X)$ is the *normal* or *Gaussian* density with mean γ and variance θ, which is given by

$$p(X) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{(X-\gamma)^2}{2\theta} \right) .$$

Let $\{X(d)\}_{d=1}^{\infty}$ be a sequence of IID random variables drawn from $p(X)$. Let $\{Y(d)\}_{d=1}^{\infty}$ be the sequence of random variables defined by

$$Y(d) = \frac{1}{d} \sum_{d'=1}^{d} X(d') \qquad \text{for every } d = 1, \cdots, \infty .$$

Normal Growth Rate Model

We can easily check that

$$\mathrm{Ex}(Y(d)) = \gamma \,, \qquad \mathrm{Var}(Y(d)) = \frac{\theta}{d} \,.$$

We can also check that

$$\mathrm{Ex}(Y(d)|Y(d-1)) = \frac{d-1}{d} Y(d-1) + \frac{1}{d}\gamma \,.$$

So the variables $Y(d)$ are neither independent nor identically distributed.

It can be shown (the details are not given here) that $Y(d)$ is drawn from the normal density with mean γ and variance θ/d, which is given by

$$p_d(Y) = \sqrt{\frac{d}{2\pi\theta}} \exp\left(-\frac{(Y-\gamma)^2 d}{2\theta}\right) \,.$$

Normal Growth Rate Model

Because $S(d)/s(0) = e^{d\,Y(d)}$, the mean return at day d is

$$\begin{aligned}\mathrm{Ex}\Big(e^{d\,Y(d)}\Big) &= \sqrt{\frac{d}{2\pi\theta}}\int \exp\left(-\frac{(Y-\gamma)^2 d}{2\theta} + d\,Y\right)\mathrm{d}Y \\ &= \sqrt{\frac{d}{2\pi\theta}}\int \exp\left(-\frac{(Y-\gamma-\theta)^2 d}{2\theta} + d(\gamma+\tfrac{1}{2}\theta)\right)\mathrm{d}Y \\ &= \exp\Big(d(\gamma+\tfrac{1}{2}\theta)\Big)\,.\end{aligned}$$

Because $p_d(Y)$ becomes sharply peaked around $Y=\gamma$ as d increases, most investors will see the lower growth rate γ rather than $\gamma+\frac{1}{2}\theta$.

By setting $d=1$ in the above formula, we see that the return mean is

$$\mu = \mathrm{Ex}(R) = \mathrm{Ex}\Big(e^X - 1\Big) = \exp(\gamma+\tfrac{1}{2}\theta) - 1\,.$$

Hence, $\mu > \gamma+\frac{1}{2}\theta$, with $\mu \approx \gamma+\frac{1}{2}\theta$ when $(\gamma+\frac{1}{2}\theta) << 1$.

Normal Growth Rate Model

Therefore most investors will see a return that is below the return mean μ — far below in volatile markets. This is because e^X amplifies the tail of the normal density. For a more realistic IID model with a density $p(X)$ that decays more slowly than a normal density as $X \to \infty$, this difference can be more striking. Said another way, most investors will not see the same return as Warren Buffett, but his return will boost the mean.

The normal growth rate model confirms that γ is a better proxy for how well a risky asset might perform than μ because $p_d(Y)$ becomes more peaked around $Y = \gamma$ as d increases. We will extend this result to a general class of IID models that are more realistic.

Independent, Identically-Distributed Models for Markets

We now consider a market with N risky assets. Let $\{s_i(d)\}_{d=0}^D$ be the share price history of asset i. The associated return and growth rate histories are $\{r_i(d)\}_{d=1}^D$ and $\{x_i(d)\}_{d=1}^D$ where

$$r_i(d) = \frac{s_i(d)}{s_i(d-1)} - 1\,, \qquad x_i(d) = \log\left(\frac{s_i(d)}{s_i(d-1)}\right)\,.$$

Because each $s_i(d)$ is positive, each $r_i(d)$ is in $(-1, \infty)$, and each $x_i(d)$ is in $(-\infty, \infty)$. Let $\mathbf{r}(d)$ and $\mathbf{x}(d)$ be the N-vectors

$$\mathbf{r}(d) = \begin{pmatrix} r_1(d) \\ \vdots \\ r_N(d) \end{pmatrix}\,, \qquad \mathbf{x}(d) = \begin{pmatrix} x_1(d) \\ \vdots \\ x_N(d) \end{pmatrix}\,.$$

The market return and growth rate histories can then be expressed simply as $\{\mathbf{r}(d)\}_{d=1}^D$ and $\{\mathbf{x}(d)\}_{d=1}^D$ respectively.

Independent, Identically-Distributed Models for Markets

An IID model for this market draws D random vectors $\{\mathbf{R}(d)\}_{d=1}^{D}$ from a fixed probablity density $q(\mathbf{R})$ over $(-1, \infty)^N$. Such a model is reasonable when the points $\{(d, \mathbf{r}(d))\}_{d=1}^{D}$ are distributed uniformly in d. This is hard to visualize when N is not small.

You might think a necessary condition for the entire market to have an IID model is that each asset has an IID model. *This can be visualized for each asset by plotting the points $\{(d, r_i(d))\}_{d=1}^{D}$ in the dr-plane and seeing if they appear to be distributed uniformly in d.*

Similar visual tests based on pairs of assets can be carried out by plotting the points $\{(d, r_i(d), r_j(d))\}_{d=1}^{D}$ in $\mathbb{R}^3$ with an interactive 3D graphics package.

Independent, Identically-Distributed Models for Markets

Visual tests like those described above often show that funds behave more like IID models than individual stocks or bonds. This means that portfolio balancing strategies based on IID models might work better for portfolios composed largely of funds. This is one reason why some investors prefer investing in funds over investing in individual stocks and bonds.

A better lesson to be drawn from the observation in the last paragraph is that every sufficiently diverse portfolio of assets in a market will behave more like an IID model than many of the individual assets in that market. In other words, IID models for a market can be used to develop portfolio balancing strategies when the portfolios considered are sufficiently diverse, even when the behavior of individual assets in that market may not be well described by the model. This is another reason to prefer holding diverse, broad-based portfolios.

Independent, Identically-Distributed Models for Markets

More importantly, this suggests that it is better to apply visual tests like those described above to representative portfolios rather than to individual assets in the market.

Remark. Such visual tests can only warn you when IID models might not be appropriate for describing the data. There are also statistical tests that can play this role. *There is no visual or statistical test that can insure the validity of using an IID model for a market. However, due to their simplicity, IID models are often used unless there is a good reason not to use them.*

Independent, Identically-Distributed Models for Markets

After we have decided to use an IID model for the market, we must gather statistical information about the return probability density $q(\mathbf{R})$. The mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Xi}$ of $\mathbf{R}$ are given by

$$\boldsymbol{\mu} = \int \mathbf{R}\, q(\mathbf{R})\, \mathrm{d}\mathbf{R}\,, \qquad \boldsymbol{\Xi} = \int (\mathbf{R} - \boldsymbol{\mu})(\mathbf{R} - \boldsymbol{\mu})^{\mathrm{T}} q(\mathbf{R})\, \mathrm{d}\mathbf{R}\,.$$

Given any sample $\{\mathbf{R}(d)\}_{d=1}^{D}$ drawn from $q(\mathbf{R})$, these have the unbiased estimators

$$\hat{\boldsymbol{\mu}} = \sum_{d=1}^{D} w(d)\, \mathbf{R}(d)\,, \qquad \hat{\boldsymbol{\Xi}} = \sum_{d=1}^{D} \frac{w(d)}{1 - \bar{w}} \left(\mathbf{R}(d) - \hat{\boldsymbol{\mu}}\right)\left(\mathbf{R}(d) - \hat{\boldsymbol{\mu}}\right)^{\mathrm{T}}\,.$$

If we assume that such a sample is given by the return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ then these estimators are given in terms of the vector $\mathbf{m}$ and matrix $\mathbf{V}$ by

$$\hat{\boldsymbol{\mu}} = \mathbf{m}\,, \qquad \hat{\boldsymbol{\Xi}} = \frac{1}{1 - \bar{w}}\, \mathbf{V}\,.$$

Independent, Identically-Distributed Models for Portfolios

Recall that the value of a portfolio that holds a risk-free balance $b_{\mathrm{rf}}(d)$ with return μ_{rf} and $n_i(d)$ shares of asset i during trading day d is

$$\pi(d) = b_{\mathrm{rf}}(d)\,(1+\mu_{\mathrm{rf}}) + \sum_{i=1}^{N} n_i(d) s_i(d)\,.$$

We will assume that $\pi(d) > 0$ for every d. Then the return $r(d)$ and growth rate $x(d)$ for this portfolio on trading day d are given by

$$r(d) = \frac{\pi(d)}{\pi(d-1)} - 1\,, \qquad x(d) = \log\left(\frac{\pi(d)}{\pi(d-1)}\right)\,.$$

Recall that the return $r(d)$ for the Markowitz portfolio with allocation $\mathbf{f}$ can be expressed in terms of the vector $\mathbf{r}(d)$ as

$$r(d) = (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{f}^{\mathrm{T}}\mathbf{r}(d)\,.$$

Independent, Identically-Distributed Models for Portfolios

This implies that if the underlying market has an IID model with return probability density $q(\mathbf{R})$ then the Markowitz portfolio with allocation $\mathbf{f}$ has the IID model with return probability density $q_{\mathbf{f}}(R)$ given by

$$q_{\mathbf{f}}(R) = \int \delta\Big(R - (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} - \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\; q(\mathbf{R})\,\mathrm{d}\mathbf{R}\,.$$

Here $\delta(\,\cdot\,)$ denotes the *Dirac delta distribution*, which can be defined by the property that for every sufficiently nice function $\psi(R)$

$$\int \psi(R)\,\delta\Big(R - (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} - \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\,\mathrm{d}R = \psi\Big((1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\,.$$

Independent, Identically-Distributed Models for Portfolios

Hence, by combining the foregoing formula for $q_{\mathbf{f}}(R)$ with the defining property of the Dirac delta distribution, we see that for every sufficiently nice function $\psi(R)$ we have the formula

$$\begin{aligned}
\mathrm{Ex}\big(\psi(R)\big) &= \int \psi(R)\, q_{\mathbf{f}}(R)\, \mathrm{d}R \\
&= \int \psi(R) \left[\int \delta\Big(R - (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} - \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\, q(\mathbf{R})\, \mathrm{d}\mathbf{R} \right] \mathrm{d}R \\
&= \int \left[\psi(R)\, \delta\Big(R - (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} - \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\, \mathrm{d}R \right] q(\mathbf{R})\, \mathrm{d}\mathbf{R} \\
&= \int \psi\Big((1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\, q(\mathbf{R})\, \mathrm{d}\mathbf{R}\,.
\end{aligned}$$

This formula can also be viewed as defining $q_{\mathbf{f}}(r)$.

Independent, Identically-Distributed Models for Portfolios

In particular, we can compute the mean μ of $q_{\mathbf{f}}(R)$ as

$$\begin{aligned}\mu = \mathrm{Ex}(R) &= \int \left((1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f}\right) q(\mathbf{R})\,\mathrm{d}\mathbf{R} \\ &= (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} \int q(\mathbf{R})\,\mathrm{d}\mathbf{R} + \left(\int \mathbf{R}\, q(\mathbf{R})\,\mathrm{d}\mathbf{R}\right)^{\mathrm{T}} \mathbf{f} \\ &= (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{f}\,,\end{aligned}$$

where in the last step we have used the facts that

$$\int q(\mathbf{R})\,\mathrm{d}\mathbf{R} = 1\,, \qquad \int \mathbf{R}\, q(\mathbf{R})\,\mathrm{d}\mathbf{R} = \boldsymbol{\mu}\,.$$

Independent, Identically-Distributed Models for Portfolios

This formula for μ can then be used to compute the variance ξ of $q_{\mathbf{f}}(R)$ as

$$\begin{aligned}\xi = \mathrm{Ex}\big((R-\mu)^2\big) &= \int \big((1-\mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f} - \mu\big)^2 q(\mathbf{R})\,\mathrm{d}\mathbf{R} \\ &= \int \left(\mathbf{R}^{\mathrm{T}}\mathbf{f} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{f}\right)^2 q(\mathbf{R})\,\mathrm{d}\mathbf{R} \\ &= \int \mathbf{f}^{\mathrm{T}}(\mathbf{R}-\boldsymbol{\mu})\,(\mathbf{R}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{f}\,q(\mathbf{R})\,\mathrm{d}\mathbf{R} \\ &= \mathbf{f}^{\mathrm{T}}\left(\int (\mathbf{R}-\boldsymbol{\mu})(\mathbf{R}-\boldsymbol{\mu})^{\mathrm{T}} q(\mathbf{R})\,\mathrm{d}\mathbf{R}\right)\mathbf{f} = \mathbf{f}^{\mathrm{T}}\boldsymbol{\Xi}\,\mathbf{f}\,,\end{aligned}$$

where in the last step we have used the fact that

$$\int (\mathbf{R}-\boldsymbol{\mu})(\mathbf{R}-\boldsymbol{\mu})^{\mathrm{T}} q(\mathbf{R})\,\mathrm{d}\mathbf{R} = \boldsymbol{\Xi}\,.$$

Independent, Identically-Distributed Models for Portfolios

If we assume that the return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ is an IID sample drawn from a probability density $q(\mathbf{R})$ then unbiased estimators of the associated mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Xi}$ are given in terms of $\mathbf{m}$ and $\mathbf{V}$ by

$$\hat{\boldsymbol{\mu}} = \mathbf{m}\,, \qquad \hat{\boldsymbol{\Xi}} = \frac{1}{1-\bar{w}}\,\mathbf{V}\,.$$

Moreover, the Markowitz portfolio with allocation $\mathbf{f}$ has the return history $\{r(d)\}_{d=1}^{D}$ where

$$r(d) = (1 - \mathbf{1}^{\mathrm{T}}\mathbf{f})\mu_{\mathrm{rf}} + \mathbf{f}^{\mathrm{T}}\mathbf{r}(d)\,.$$

This return history is an IID sample drawn from the probability density $q_{\mathbf{f}}(R)$ and the formulas on the last two pages show that the mean μ and variance ξ of $q_{\mathbf{f}}(R)$ have the unbiased estimators

$$\hat{\mu} = \mu_{\mathrm{rf}}(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}) + \mathbf{m}^{\mathrm{T}}\mathbf{f}\,, \qquad \hat{\xi} = \frac{1}{1-\bar{w}}\,\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,.$$