Amath 546/Econ 589

Estimating Risk Measures and Risk Budgets

Eric Zivot

April 17, 2013

Outline

- Estimating Risk Measures from Return Data
- Nonparametric Estimates
- Parametric Estimates
- Simulation-based Estimates
- Estimating Portfolio Risk Budgets from Return Data

Reading

- FRF chapter 5, sections 2 and 3
- QRM chapter 2, section 3
- FMUND chapter 5, sections 2-5; chapter 8
- SADFE chapter 5 and chapter 18

Estimating Asset Risk Measures from Return Data

Let R denote the simple return on an asset or portfolio with CDF F_R and pdf f_R. Let $RM = \sigma, VaR_\alpha$ and ES_α be the risk measures of interest. Given a confidence level α and initial investment V_0 the loss-based risk measures are

$$\begin{aligned}
\sigma_L &= V_0 \left(E[(R-\mu_R)^2] \right)^{1/2} = V_0 \sigma_R \\
VaR_\alpha &= -V_0 q^R_{1-\alpha} = -V_0 F_R^{-1}(1-\alpha) \\
ES_\alpha &= -V_0 E[R | R \leq q^R_{1-\alpha}] = \frac{V_0}{1-\alpha} \int_{-\infty}^{q^R_{1-\alpha}} x f_R(x) dx
\end{aligned}$$

We are interested in estimating

$$\sigma_R,\ q^R_{1-\alpha} \text{ and } E[R | R \leq q^R_{1-\alpha}]$$

from an observed sample of returns $\{R_1 = r_1, \ldots, R_T = r_T\}$.

Estimation Approaches

- Nonparametric
  - Use empirical distribution to estimate risk measures
- Parametric
  - Specify parametric distribution for returns,
  - estimate parameters of distribution,
  - estimate risk measures as functions of estimated parameters
- Semi-parametric (Cornish-Fisher)

Estimation Error

- Risk measure estimates are subject to estimation error

- Standard errors and confidence intervals can be used to gauge magnitude of estimation error

- Bootstrapping is often the easiest way to compute standard errors and confidence intervals

- Estimation error is often ignored in practice!

Assumptions

- Assume $R_t \sim iid\ F_R$ for $t = 1, \ldots, T$ where F_R denotes the probability distribution of R

- $E[R_t] = \mu_R < \infty$

- $var(R_t) = \sigma_R^2 < \infty$

- $cov(R_t, R_s) = 0$ for all $t \neq s$

Note: Here, risk measures are based on the unconditional distribution of returns

Nonparametric Estimation of Risk Measures

Idea: F_R is unknown and you estimate F_R using the empirical distribution

$$\hat{F}_R(r) = \frac{\#\text{ of returns less than or equal to } r}{T} = \frac{\sum_{t=1}^{T} 1\{R_t \leq r\}}{T}$$

$$1\{R_t \leq r\} = 1 \text{ if } R_t \leq r;\ 0 \text{ otherwise}$$

Properties

- $\hat{F}_R(r) \rightarrow F_R(r)$ as $T \rightarrow \infty$ where $F_R(r) = \Pr(R \leq r) = \int_{-\infty}^{r} f_R(x)dx$

Return Volatility σ

$$\hat{\sigma}_R = \left(\frac{1}{T-1} \sum_{t=1}^{T} (R_t - \hat{\mu}_R)^2 \right)^{1/2}, \ \hat{\mu}_R = \frac{1}{T} \sum_{t=1}^{T} R_t$$

Properties

$$E[\hat{\sigma}_R] \neq \sigma_R$$

$$\sigma_R \xrightarrow{p} \sigma_R \text{ as } T \rightarrow \infty$$

$$\hat{\sigma}_R \overset{A}{\sim} N\left(\sigma_R, \frac{\sigma_R^2}{2T}\right)$$

$$SE(\hat{\sigma}_R) = \frac{\sigma_R}{\sqrt{2T}}$$

$$\widehat{SE}(\hat{\sigma}_R) = \frac{\hat{\sigma}_R}{\sqrt{2T}}$$

Return quantile $q_{1-\alpha}^{R}$

$$\begin{aligned}\hat{q}_{1-\alpha}^{R} &= \text{ empirical } 1-\alpha \text{ quantile} \\ &= r_t \text{ such that } 100 \times (1-\alpha) \text{ \% of data is less than } r_t \\ \widehat{VaR}_{\alpha}^{HS} &= -V_0 \times \hat{q}_{1-\alpha}^{R},\ HS = \text{ "historical simulation"}\end{aligned}$$

Properties

$$\hat{q}_{1-\alpha}^{R} \rightarrow q_{1-\alpha}^{R} \text{ as } T \rightarrow \infty$$

$$\hat{q}_{1-\alpha}^{R} \overset{A}{\sim} N\left(q_{1-\alpha}^{R}, \frac{\alpha(1-\alpha)}{T \times f_R(q_{1-\alpha}^{R})^2}\right) \text{ which depends on } f_R$$

$$SE(\hat{q}_{1-\alpha}^{R}) = \frac{\sqrt{\alpha(1-\alpha)}}{\sqrt{T} \times f_R(q_{1-\alpha}^{R})}$$

Note: Estimating $SE(\hat{q}_{1-\alpha}^{R})$ requires estimating $f_R(q_{1-\alpha}^{R})$

Return tail average $E[R|R \leq q^R_{1-\alpha}]$

$$\begin{aligned}
\widehat{E}[R|R &\leq& q^R_{1-\alpha}] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^{T} R_t \cdot 1\left\{R_t \leq \hat{q}^R_{1-\alpha}\right\} \\
1\left\{R_t \leq \hat{q}^R_{\alpha}\right\} &=& 1 \text{ if } R_t \leq \hat{q}^R_{1-\alpha};\ 0 \text{ otherwise} \\
B_{1-\alpha} &=& \sum_{t=1}^{T} 1\left\{R_t \leq \hat{q}^R_{1-\alpha}\right\} = \#\text{ of returns } \leq \hat{q}^R_{1-\alpha} \\
\widehat{ES}^{HS}_{\alpha} &=& -V_0 \times \widehat{E}[R|R \leq \hat{q}^R_{1-\alpha}]
\end{aligned}$$

Properties

$$\begin{aligned}
\widehat{E}[R|R &\leq& q^R_{1-\alpha}] \rightarrow E[R|R \leq q^R_{1-\alpha}] \text{ as } T \rightarrow \infty \\
\widehat{E}[R|R &\leq& q^R_{1-\alpha}] \overset{A}{\sim} N\left(E[R|R \leq q^R_{1-\alpha}], \frac{\sigma^2_R}{B_{1-\alpha}}\right)
\end{aligned}$$

Parametric Estimation of Risk Measures

- Assume $F_R = F_R(r; \boldsymbol{\theta})$ and $f_R(r; \boldsymbol{\theta})$ are parametric CDFs and pdfs that depend on p unknown parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)'$.

- Estimate $\boldsymbol{\theta}$ (typically by maximum likelihood) from observed data giving $\hat{\boldsymbol{\theta}}$, $F_R(r, \hat{\boldsymbol{\theta}})$ and $f_R(r; \hat{\boldsymbol{\theta}})$

- Estimate risk measures from $F_R(r, \hat{\boldsymbol{\theta}})$ and $f_R(r; \hat{\boldsymbol{\theta}})$.

Maximum Likelihood Estimation of θ

Let $R_1, \ldots, R_n$ be an iid sample with pdf $f_R(r_i; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a $(p \times 1)$ vector of parameters

The *joint density* of the sample $\mathbf{r} = (r_1, \ldots, r_n)'$ is, by independence, equal to the product of the marginal densities

$$f_R(\mathbf{r}; \boldsymbol{\theta}) = f_R(r_1; \boldsymbol{\theta}) \cdots f_R(r_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f_R(r_i; \boldsymbol{\theta}).$$

The *likelihood function* is defined as the joint density treated as a functions of the parameters $\boldsymbol{\theta}$:

$$L_R(\boldsymbol{\theta}|\mathbf{r}) = f_R(\mathbf{r}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f_R(r_i; \boldsymbol{\theta}).$$

The *maximum likelihood estimator*, denoted $\hat{\boldsymbol{\theta}}_{mle}$, is the value of $\boldsymbol{\theta}$ that maximizes $L_R(\boldsymbol{\theta}|\mathbf{r})$. That is,

$$\hat{\boldsymbol{\theta}}_{mle} = \arg\max_{\boldsymbol{\theta}} L_R(\boldsymbol{\theta}|\mathbf{r})$$

It usually much easier to maximize the log-likelihood function $\ln L_R(\boldsymbol{\theta}|\mathbf{r})$. Since $\ln(\cdot)$ is a monotonic function, equivalently

$$\hat{\boldsymbol{\theta}}_{mle} = \arg\max_{\boldsymbol{\theta}} \ln L_R(\boldsymbol{\theta}|\mathbf{r})$$

With random sampling, the log-likelihood has the particularly simple form

$$\ln L_R(\boldsymbol{\theta}|\mathbf{r}) = \ln\left(\prod_{i=1}^{n} f_R(r_i;\boldsymbol{\theta})\right) = \sum_{i=1}^{n} \ln f_R(r_i;\boldsymbol{\theta})$$

Typically, we find the MLE by differentiating $\ln L_R(\boldsymbol{\theta}|\mathbf{r})$ and solving the first order conditions

$$\frac{\partial \ln L_R(\hat{\theta}_{mle}|\mathbf{r})}{\partial \boldsymbol{\theta}} = \mathbf{0}$$

Since $\boldsymbol{\theta}$ is $(p \times 1)$ the first order conditions define p, potentially nonlinear, equations in p unknown values:

$$\frac{\partial \ln L_R(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial \ln L_R(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r})}{\partial \theta_1} \\ \vdots \\ \frac{\partial \ln L_R(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r})}{\partial \theta_p} \end{pmatrix} = \mathbf{0}$$

The $p \times 1$ vector of derivatives of the log-likelihood function is called the *score* vector

$$S_R(\boldsymbol{\theta}|\mathbf{r}) = \frac{\partial \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \boldsymbol{\theta}}$$

By definition, the MLE satisfies

$$S_R(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r}) = \mathbf{0}$$

The $p \times p$ matrix of second derivatives of the log-likelihood is called the *Hessian*

$$H_R(\boldsymbol{\theta}|\mathbf{r}) = \frac{\partial^2 \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \frac{\partial^2 \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \theta_1^2} & \cdots & \frac{\partial^2 \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \theta_1 \partial \theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \theta_p \partial \theta_1} & \cdots & \frac{\partial^2 \ln L_R(\boldsymbol{\theta}|\mathbf{r})}{\partial \theta_p^2} \end{pmatrix}$$

The *information matrix* is defined as minus the expectation of the Hessian

$$I_R(\boldsymbol{\theta}|\mathbf{r}) = -E[H_R(\boldsymbol{\theta}|\mathbf{r})]$$

Computing the MLE

- With certain simple pdfs $f_R(r_i;\boldsymbol{\theta})$ (e.g., normal pdf) $\hat{\boldsymbol{\theta}}_{mle}$ can be obtained in closed form.

- In general, numerical maximization of $\ln L_R(\boldsymbol{\theta}|\mathbf{r})$ is based on Newton-Raphson iteration

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{mle,n+1} &= \hat{\boldsymbol{\theta}}_{mle,n} - H(\hat{\boldsymbol{\theta}}_{mle,n}|\mathbf{r})^{-1}S(\hat{\boldsymbol{\theta}}_{mle,n}|\mathbf{r}) \\ \hat{\boldsymbol{\theta}}_{mle,0} &= \boldsymbol{\theta}_0 = \text{starting values}\end{aligned}$$

  and iteration stops when $S(\hat{\boldsymbol{\theta}}_{mle,n}|\mathbf{r}) \approx \mathbf{0}$.

Asymptotic Properties of Maximum Likelihood Estimators

Under general regularity conditions, $\hat{\boldsymbol{\theta}}_{mle}$ has the following asymptotic properties as $T \to \infty$

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_{mle} &\xrightarrow{p} \theta \\
\hat{\boldsymbol{\theta}}_{mle} &\sim N\left(\boldsymbol{\theta}, \frac{1}{T} H(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r})^{-1}\right) \\
SE(\hat{\theta}_{i,mle}) &= \text{ ith diagonal element of } \frac{1}{T} H(\hat{\boldsymbol{\theta}}_{mle}|\mathbf{r})^{-1}
\end{aligned}
$$

Parametric Estimates of VaR_α and ES_α

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \hat{\boldsymbol{\theta}}_{mle} \\
F_R(r, \hat{\boldsymbol{\theta}}) &= \text{parametric estimate of } F_R \\
F_R^{-1}(r, \hat{\boldsymbol{\theta}}) &= \text{parametric estimate of } F_R^{-1} \\
f_R(r, \hat{\boldsymbol{\theta}}) &= \text{ parametric estimate of } f_R
\end{aligned}$$

Then

$$\begin{aligned}
\widehat{VaR}_\alpha^{par} &= -V_0 \times F_R^{-1}(1-\alpha; \hat{\boldsymbol{\theta}}) = -V_0 q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}) \\
\widehat{ES}_\alpha^{par} &= \frac{-V_0}{1-\alpha} \int_{-\infty}^{q_{1-\alpha}^R(\hat{\boldsymbol{\theta}})} x f_R(x; \hat{\boldsymbol{\theta}}) dx
\end{aligned}$$

Here, the superscript "par" refers to "parametric distribution".

Example: Normal Distribution

$$R_t \sim iid\ N(\mu_R, \sigma_R^2),\ \boldsymbol{\theta} = (\mu_R, \sigma_R^2)'$$

$$\mathbf{r} = (r_1, \ldots, r_T) = \text{ observed sample}$$

$$f_R(r; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{r-\mu_r}{\sigma_r}\right)^2\right),$$

$$\ln L(\boldsymbol{\theta}|\mathbf{r}) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma_r^2) - \frac{1}{2\sigma_r^2}\sum_{t=1}^{T}(r_t - \mu_r)^2$$

$$\hat{\mu}_{R,mle} = \frac{1}{T}\sum_{t=1}^{T} r_t,\ \hat{\sigma}^2_{R,mle} = \frac{1}{T}\sum_{t=1}^{T}\left(r_t - \hat{\mu}_{R,mle}\right)^2$$

$$\hat{\boldsymbol{\theta}}_{mle} = (\hat{\mu}_{R,mle}, \hat{\sigma}^2_{R,mle})'$$

Example: Normal Distribution continued

$$\begin{aligned}
F_R^{-1}(1-\alpha;\hat{\boldsymbol{\theta}}_{mle}) &= q_{1-\alpha}^{R}(\hat{\boldsymbol{\theta}}_{mle}) = \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times q_{1-\alpha}^{Z} \\
q_{1-\alpha}^{Z} &= 1-\alpha \text{ quantile of } Z \sim N(0,1) \\
E[R|R &\leq q_{1-\alpha}^{R}(\hat{\boldsymbol{\theta}}_{mle})] = -\left(\hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times \frac{\phi(q_{1-\alpha}^{Z})}{1-\alpha}\right) \\
\widehat{VaR}_{\alpha}^{norm} &= -V_0 \times \left(\hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times q_{1-\alpha}^{Z}\right) \\
\widehat{ES}_{\alpha}^{norm} &= -V_0 \left(\hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times \frac{\phi(q_{1-\alpha}^{Z})}{1-\alpha}\right)
\end{aligned}$$

Accuracy of VaR and ES Estimates

Using the Central Limit Theorem Result (see homework)

$$\begin{pmatrix} \hat{\mu}_{R,mle} \\ \hat{\sigma}_{R,mle} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_R \\ \sigma_R \end{pmatrix}, \begin{pmatrix} \frac{\sigma_R^2}{T} & 0 \\ 0 & \frac{\sigma_R^2}{2T} \end{pmatrix} \right)$$

it is straightforward to show that

$$SE\left(q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}_{mle})\right) = \frac{\sigma_R}{\sqrt{T}}\left(1 + \frac{1}{2}\left(q_{1-\alpha}^Z\right)^2\right)^{1/2}$$

For a given values of σ_R and T, $SE\left(q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}_{mle})\right)$ increases nonlinearly with α. Very small quantiles (i.e., $1 - \alpha \approx 0$) are estimated very imprecisely.

Example: Student's t Distribution (see QRM pgs 40 and 46)

Recall, a standardized Student's t random variable is defined as

$$\begin{aligned} t &= \frac{Z}{\sqrt{\frac{X}{v}}},\ Z \sim N(0,1),\ X \sim \chi^2_v \\ & X \text{ and } Z \text{ are independent} \\ E[t] &= 0,\ var(t) = \frac{v}{v-2} \text{ for } v > 2 \end{aligned}$$

A Student's t random variable with mean μ and scale σ^2 (not variance) is defined as

$$\begin{aligned} Y &= \mu + \sigma \times t \\ E[Y] &= \mu,\ var(Y) = \frac{\sigma^2 v}{v-2} \neq \sigma^2 \text{ for } v > 2 \end{aligned}$$

Assume

$$
\begin{aligned}
& R_t \sim iid\ t(\mu_R, \sigma_R^2, v_R),\ \ \boldsymbol{\theta} = (\mu_R, \sigma_R^2, v_R)' \\
E[R_t] &= \mu_R,\ var(R_t) = \frac{\sigma_R^2 v_R}{(v_R - 2)} \neq \sigma_R^2 \\
\mathbf{r} &= (r_1, \ldots, r_T) = \text{ observed sample} \\
f_R(r; \boldsymbol{\theta}) &= \left[\frac{\Gamma((v_R + 1)/2}{(\sigma_R^2 \pi v_R)^{1/2} \Gamma(v_R/2)}\right] \left(1 + \frac{1}{v_R}\left(\frac{r - \mu_R}{\sigma_R}\right)^2\right)^{-(v_R+1)/2} \\
\ln L(\boldsymbol{\theta}|\mathbf{r}) &= \sum_{i=1}^{n} \ln f_R(r_i; \boldsymbol{\theta}) = \text{complicated nonlinear equation in } \boldsymbol{\theta}
\end{aligned}
$$

Here, there are no closed form solutions for $\hat{\mu}_{R.mle}$, $\hat{\sigma}^2_{R,mle}$ and $\hat{v}_{R,mle}$. We have to obtain these values by numerical maximization of the log-likelihood.

Example: Student's t Distribution continued

$$F_R^{-1}(1-\alpha;\hat{\boldsymbol{\theta}}_{mle}) = q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}_{mle}) = \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times q_{1-\alpha}^{t_{\hat{v}_{mle}}}$$

$$q_{1-\alpha}^{t_{\hat{v}_{mle}}} = 1-\alpha \text{ quantile of standard Student's t with } \hat{v}_{R,mle} \text{ df}$$

$$E[R|R \leq q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}_{mle})] = -(\hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle})$$

$$\times \left(\frac{f_{\hat{v}_{mle}}(q_{1-\alpha}^{t_{\hat{v}_{mle}}})}{1-\alpha} \left(\frac{\hat{v}_{mle} + \left(q_{1-\alpha}^{t_{\hat{v}_{mle}}}\right)^2}{\hat{v}_{mle}-1} \right) \right)$$

$$\widehat{VaR}_\alpha^t = -V_0 \times \left(\hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \times q_{1-\alpha}^{t_{\hat{v}_{mle}}}\right)$$

$$\widehat{ES}_\alpha^t = -V_0 \times E[R|R \leq q_{1-\alpha}^R(\hat{\boldsymbol{\theta}}_{mle})]$$

Remark

- Suppose you don't know how to calculate $E[R|R \leq q^R_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle})]$ when $R \sim t(\mu_R, \sigma_R, v_R)$

- You could easily approximate $E[R|R \leq q^R_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle})]$ by Monte Carlo simulation

  – Simulate N values $\tilde{R}_1, \ldots, \tilde{R}_N$ from $f(r, \hat{\boldsymbol{\theta}}_{mle})$

  – Approximate $E[R|R \leq q^R_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle})]$ using nonparametric estimate

$$\hat{E}[R|R \leq q^R_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle})] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^{T} \tilde{R}_t \cdot 1\left\{\tilde{R}_t \leq \hat{q}^{\tilde{R}}_{1-\alpha}\right\}$$

Semi-Parametric Estimation

- Hybrid of nonparametric and parametric estimation
  - Some parts of F_R are treated nonparametrically and some parts are treated parametrically

Example: Cornish-Fisher (Modified) VaR

Idea: Approximate unknown CDF F_Z of $Z = (R - \mu_R)/\sigma_R$ using 2 term Edgeworth expansion around normal CDF $\Phi(\cdot)$ and invert expansion to get quantile approximation:

$$\begin{aligned} F_{Z,CF}^{-1}(1-\alpha) &= z_{1-\alpha} + \frac{1}{6}(z_{1-\alpha}^2 - 1) \times skew + \frac{1}{24}(z_{1-\alpha}^3 - 3z_{1-\alpha}) \times ekurt \\ &\quad - \frac{1}{36}(2z_{1-\alpha}^3 - 5z_{1-\alpha}) \times skew \\ z_{1-\alpha} &= \Phi^{-1}(1-\alpha) = N(0,1) \text{ quantile}, \\ skew &= E[Z^3],\ ekurt = E[Z^4] - 3 \end{aligned}$$

This quantile approximation is called the Cornish-Fisher approximation. The Cornish-Fisher return quantile and VaR are

$$\begin{aligned} q_{1-\alpha}^{CF} &= \mu_R + \sigma_R \times F_{Z,CF}^{-1}(1-\alpha) \\ VaR_\alpha^{CF} &= -V_0 \times q_{1-\alpha}^{CF} \end{aligned}$$

Remarks:

- The values of $skew = E[Z^3]$ and $ekurt = E[Z^4] - 3$ depend on the unknown CDF F_Z and are estimated nonparametrically using sample estimates $\widehat{skew}$ and $\widehat{ekurt}$ computed from returns. The estimated CF quantile is

$$\hat{q}_{1-\alpha}^{CF} = \hat{\mu}_R + \hat{\sigma}_R \times \hat{F}_{Z,CF}^{-1}(1-\alpha)$$

- Because $\hat{q}_{1-\alpha}^{CF}$ is an approximation based on sample estimates, it can produce strange results in some cases and should be used with care.

- For modified ES, See Boudt, Peterson and Croux (2008) "Estimation and Decomposition of Downside Risk for Portfolios with Nonnormal Returns," *Journal of Risk.*

- The PerformanceAnalytics functions `VaR()` and `ES()` compute the Cornish-Fisher (aka modified) VaR and ES quantities

Estimating Portfolio Risk Measures and Risk Budgets

Let $\mathbf{R} = (R_1, \ldots, R_n)'$ denote the vector of simple returns on n assets, and let $\mathbf{w} = (w_1, \ldots, w_n)'$ denote portfolio weights such that $\sum_{i=1}^{n} w_i = 1$.

Assumptions

- $\mathbf{R}_t = (R_{1t}, \ldots, R_{nt})'$ is iid with joint CDF $F_{\mathbf{R}}$ and pdf $f_{\mathbf{R}}$
- $E[\mathbf{R}_t] = \boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ for all t
- $var(\mathbf{R}_t) = E[(\mathbf{R}_t - \boldsymbol{\mu})(\mathbf{R}_t - \boldsymbol{\mu})'] = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{pmatrix}$

Portfolio Return Distribution

- $R_{pt} = \mathbf{w}'\mathbf{R}_t = \sum_{i=1}^{N} w_i R_{it}$
  - $R_{pt} \sim iid\ F_{R_p}$ which depends on the joint distribution $F_{\mathbf{R}}$
- $\mu_p = \mathbf{w}'\boldsymbol{\mu}$,
- $\sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$ and $\sigma_p = (\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w})^{1/2}$

Fixed and Active Portfolios

Fixed Portfolio

- $\mathbf{w} = (w_1, \ldots, w_n)'$ is fixed over time (e.g., 60% stocks and 40% bonds)

Active Portfolio

- $\mathbf{w}_t = (w_{1t}, \ldots, w_{nt})'$ depends on time (e.g., portfolio manager actively rebalances portfolio every period)

Remark: Treatment of $\mathbf{w}$ influences how we compute portfolio risk measures. It what follows, we will treat $\mathbf{w}$ as fixed.

Portfolio Risk Measures

Given a confidence level α and initial investment V_0 the loss-based portfolio risk measures are

$$\begin{aligned}
\sigma_L &= V_0 \left(E[(R_p - \mu_p)^2] \right)^{1/2} = V_0 \sigma_p \\
VaR_\alpha &= -V_0 q_{1-\alpha}^{R_p} = -V_0 F_{R_p}^{-1}(1-\alpha) \\
ES_\alpha &= -V_0 E[R_p | R_p \leq q_{1-\alpha}^{R_p}] = \frac{-V_0}{1-\alpha} \int_{-\infty}^{q_{1-\alpha}^{R_p}} x f_{R_p}(x) dx
\end{aligned}$$

Note: Because $R_p = \mathbf{w}'\mathbf{R}$ the above risk measures are functions of $\mathbf{w}$

$$\sigma_L = \sigma_L(\mathbf{w}),\ VaR_\alpha = VaR_\alpha(\mathbf{w}) \text{ and } ES_\alpha = ES_\alpha(\mathbf{w})$$

Portfolio Risk Budgets

Let $RM(\mathbf{w})$ denote the risk measures $\sigma_L(\mathbf{w})$, $VaR_\alpha(\mathbf{w})$ and $ES_\alpha(\mathbf{w})$ as functions of the portfolio weights $\mathbf{w}$. The portfolio risk budgets are the quantities

$$
\begin{aligned}
MCR_i^j &= \frac{\partial RM(\mathbf{w})}{\partial w_i} = \text{asset } i \text{ marginal contribution to risk} \\
CR_i^j &= w_i \frac{\partial RM(\mathbf{w})}{\partial w_i} = \text{ asset } i \text{ contribution to risk} \\
PCR_i^j &= \frac{w_i \frac{\partial RM(\mathbf{w})}{\partial w_i}}{RM(\mathbf{w})} = \text{ asset } i \text{ percent contribution to risk} \\
j &= \sigma_L, VaR_\alpha \text{ and } ES_\alpha
\end{aligned}
$$

Recall,

$$\begin{aligned}
\frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} &= \left[\frac{1}{\sigma_p(\mathbf{w})}\mathbf{\Sigma w}\right]_i = MCR_i^{\sigma} \\
\frac{\partial VaR_\alpha(\mathbf{w})}{\partial w_i} &= E[R_i | R_p = VaR_\alpha(\mathbf{w})] = MCR_i^{VaR} \\
\frac{\partial ES_\alpha(\mathbf{w})}{\partial w_i} &= E[R_i | R_p \leq VaR_\alpha(\mathbf{w})] = MCR_i^{ES}
\end{aligned}$$

We are interested in estimating the portfolio return risk measures

$$
\begin{aligned}
\sigma_p,\ q_{1-\alpha}^{R_p},\ E[R_p|R_p &\leq q_{1-\alpha}^{R_p}] \text{ and } MCR_i^j \\
j &= \sigma,\ VaR \text{ and } ES
\end{aligned}
$$

from an observed sample of return vectors $\{\mathbf{R}_1 = \mathbf{r}_1, \ldots, \mathbf{R}_T = \mathbf{r}_T\}$.

Nonparametric Estimation of Portfolio Risk Measures

Portfolio Volatility $\sigma_p(\mathbf{w})$

Note: There are two equivalent ways to estimate portfolio volatility σ_p

Method 1:

- Create time series of portfolio returns $R_{pt} = \mathbf{w}'\mathbf{R}_t$

- Compute sample standard deviation of portfolio returns

$$\hat{\sigma}_p = \left(\frac{1}{T-1} \sum_{t=1}^{T} (R_{pt} - \hat{\mu}_p)^2 \right)^{1/2}$$

$$\hat{\mu}_p = \frac{1}{T} \sum_{t=1}^{T} R_{pt}$$

Method 2:

- Utilize the formula $\sigma_p(\mathbf{w}) = (\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w})^{-1/2}$

- Compute sample covariance matrix of return vector

$$\underset{(n\times n)}{\hat{\boldsymbol{\Sigma}}} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{R}_t - \hat{\boldsymbol{\mu}})(\mathbf{R}_t - \hat{\boldsymbol{\mu}})'$$

$$\underset{(n\times 1)}{\hat{\boldsymbol{\mu}}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t$$

- Compute $\hat{\sigma}_p(\mathbf{w}) = \left(\mathbf{w}'\hat{\boldsymbol{\Sigma}}\mathbf{w}\right)^{1/2}$

Estimating $MCR_i^{\sigma} = \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$

Recall

$$\frac{\partial \sigma_p(\mathbf{w})}{\partial \mathbf{w}} = \frac{1}{\sigma_p(\mathbf{w})} \mathbf{\Sigma w} = \left(\mathbf{w}' \mathbf{\Sigma w}\right)^{-1/2} \mathbf{\Sigma w}$$

Simply plug-in $\hat{\mathbf{\Sigma}}$ for $\mathbf{\Sigma}$ giving

$$\frac{\partial \widehat{\sigma_p(\mathbf{w})}}{\partial \mathbf{w}} = \left(\mathbf{w}' \hat{\mathbf{\Sigma}} \mathbf{w}\right)^{-1/2} \hat{\mathbf{\Sigma}} \mathbf{w}$$

Portfolio VaR and ES

$$
\begin{aligned}
\widehat{VaR}_{\alpha}^{HS} &= -V_0 \times \hat{q}_{1-\alpha}^{R_p},\ HS = \text{"historical simulation"} \\
\hat{q}_{1-\alpha}^{R} &= \text{empirical quantile of } R_{p,t} \\
\widehat{ES}_{\alpha}^{HS} &= -V_0 \times \widehat{E}[R_p|R_p \leq \hat{q}_{1-\alpha}^{R}] \\
\widehat{E}[R_p|R_p &\leq q_{1-\alpha}^{R}] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^{T} R_{pt} \cdot 1 \left\{ R_{pt} \leq \hat{q}_{1-\alpha}^{R_p} \right\}
\end{aligned}
$$

$RM(\mathbf{w}) = VaR_\alpha$ and ES_w

Assume the $n \times 1$ vector of returns $\mathbf{R}_t$ is iid but make no distributional assumptions:

$$\begin{aligned} \{\mathbf{R}_1, \ldots, \mathbf{R}_T\} &= \text{observed iid sample} \\ R_{p,t} &= \mathbf{w}'\mathbf{R}_t \end{aligned}$$

Estimate marginal contributions to risk using *historical simulation*

$$\begin{aligned} \hat{E}^{HS}[R_{it}|R_{p,t} &= VaR_\alpha] = \\ &\frac{1}{m}\sum_{t=1}^{T} R_{it} \cdot 1\left\{\widehat{VaR}_\alpha^{HS} - \varepsilon \leq R_{p,t} \leq \widehat{VaR}_\alpha^{HS} + \varepsilon\right\} \\ \hat{E}^{HS}[R_{it}|R_{p,t} &\leq VaR_\alpha] = \frac{1}{B_{1-\alpha}}\sum_{t=1}^{T} R_{it} \cdot 1\left\{R_{p,t} \leq \widehat{VaR}_\alpha^{HS}\right\} \end{aligned}$$

Here, $VaR_\alpha = q_{1-\alpha}^R$ and $\widehat{VaR}_\alpha^{HS} = \hat{q}_{1-\alpha}^R$ is the empirical $1-\alpha$ quantile of returns.

Remarks:

- The historical simulation estimator

$$\hat{E}^{HS}[R_{it}|R_{p,t} = VaR_{\alpha}] = \frac{1}{m}\sum_{t=1}^{T} R_{it} \cdot \mathbf{1}\left\{\widehat{VaR}_{\alpha}^{HS} - \varepsilon \leq R_{p,t} \leq \widehat{VaR}_{\alpha}^{HS} + \varepsilon\right\}$$

  is a special case of a so-called nonparametric "kernel" estimator where the kernel weight function is the rectangular kernel. See Yamai, Y. and T. Yoshiba (2002). "Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization," *Institute for Monetary and Economic Studies*, Bank of Japan.

- The estimator takes a local average of returns in a neighborhood of $R_{p,t} = VaR_{\alpha}$

- Other types of kernel weight functions can also be used. See Epperlein, E. and A. Smillie (2006). "Cracking VAR with Kernels," *Risk*.

Parametric Estimation of Portfolio Risk Measures: Multivariate Normal Distribution

$$\mathbf{R} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),\ \boldsymbol{\theta} = (\boldsymbol{\mu}, vech(\boldsymbol{\Sigma}))'$$
$$\Rightarrow R_p \sim N(\mu_p, \sigma_p^2),\ \mu_p = \mathbf{w}'\boldsymbol{\mu} \text{ and } \sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

MLEs of μ and $\boldsymbol{\Sigma}$

$$\begin{aligned} f_R(\mathbf{r}) &= (2\pi)^{-n/2}\det(\boldsymbol{\Sigma})^{-1/2}\exp\left(-\frac{1}{2}(\mathbf{r}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{r}-\boldsymbol{\mu})\right) \\ \hat{\boldsymbol{\mu}}_{mle} &= \frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t \\ \hat{\boldsymbol{\Sigma}}_{mle} &= \frac{1}{T}\sum_{t=1}^{T}(\mathbf{R}_t-\hat{\boldsymbol{\mu}})(\mathbf{R}_t-\hat{\boldsymbol{\mu}})' \end{aligned}$$

Note: $vech(\mathbf{\Sigma})$ stacks the diagonal and unique off diagonal elements of the $n \times n$ matrix $\mathbf{\Sigma}$ into a $n(n+1)/2$ vector

Example:

$$\underset{(3\times 3)}{\mathbf{\Sigma}} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

$$\underset{(6\times 1)}{vech(\mathbf{\Sigma})} = \begin{pmatrix} \sigma_1^2 \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_2^2 \\ \sigma_{23} \\ \sigma_3^2 \end{pmatrix}$$

Estimates of portfolio Risk Measures

$$
\begin{aligned}
\hat{\sigma}_{p,mle} &= (\mathbf{w}'\hat{\mathbf{\Sigma}}_{mle}\mathbf{w})^{1/2} \\
\hat{q}^{R_p}_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle}) &= \hat{\mu}_{p,mle} + \hat{\sigma}_{p,mle} \times q^{Z}_{1-\alpha} = \mathbf{w}'\hat{\boldsymbol{\mu}}_{mle} + \hat{\sigma}_{p,mle} \times q^{Z}_{1-\alpha} \\
E[R_p|R_p &\leq q^{R_p}_{1-\alpha}(\hat{\boldsymbol{\theta}}_{mle})] = -\left(\hat{\mu}_{p,mle} + \hat{\sigma}_{p,mle} \times \frac{\phi(q^{Z}_{1-\alpha})}{1-\alpha}\right)
\end{aligned}
$$

Remark:

- Analytic formulas exist for risk budgets (see homework 2)

Parametric Estimation of Portfolio Risk Measures: Multivariate Student's t Distribution

A $n \times 1$ multivariate Student's t random vector $\mathbf{Y}$ with mean vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and degrees of freedom v can be defined from

$$\begin{aligned} \mathbf{Y} &= \boldsymbol{\mu} + \sqrt{\frac{v}{W}}\mathbf{Z} \\ & \mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \\ & W \sim \chi^2_v \text{ independent of } \mathbf{Z} \end{aligned}$$

Here,

$$\begin{aligned} E[\mathbf{Y}] &= \boldsymbol{\mu} \\ var(\mathbf{Y}) &= \frac{v}{(v-2)}\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma} \end{aligned}$$

Result:

$$\mathbf{R} \sim t(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v) \text{ and } R_p = \mathbf{w}'\mathbf{R}$$
$$\Rightarrow R_p \sim t(\mu_p, \sigma_p^2, v),$$
$$\mu_p = \mathbf{w}'\boldsymbol{\mu}, \ \sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$
$$var(R_p) = \frac{v}{v-2}\sigma_p^2$$

Fitting the Multivariate t Distribution

SDAFE profile likelihood method (chapter 5)

- Make a grid of v values between upper and lower bounds (e.g. $v_{lower} = 2.1$, $v_{upper} = 6$)

- Find mle for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for each fixed v on grid and compute profile log-likelihood $\ln L(\boldsymbol{\mu}_{mle}, \boldsymbol{\Sigma}_{mle}, v)$

- Define $v_{mle} = \max_v \ln L(\boldsymbol{\mu}_{mle}, \boldsymbol{\Sigma}_{mle}, v)$

- Recompute mle for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using v_{mle}.

Estimating Risk Measures and Risk Budgets from Simulated Returns

- Generate B simulated values from $t(\hat{\boldsymbol{\mu}}_{mle}, \hat{\boldsymbol{\Sigma}}_{mle}, \hat{v}_{mlt})$ denoted $\{\tilde{\mathbf{R}}_t\}_1^B$
- Create B simulated portfolio returns $\tilde{R}_{p,t} = \mathbf{w}'\tilde{\mathbf{R}}_t$, $t = 1, \ldots, B$
- Estimate VaR and ES nonparametrically using $\{\tilde{R}_{p,t}\}_1^B$ and portfolio weights
- Estimate risk budgets nonparametrically using $\{\tilde{\mathbf{R}}_t\}_1^B$ and $\{\tilde{R}_{p,t}\}_1^B$

General Multivariate Distributions (some to be covered in more detail later)

- Elliptical Distributions (e.g., multivariate normal and multivariate t)

- Multivariate Skew Normal and Skew t

- Multivariate Generalized Hyperbolic

- Copula Generated Distributions (most general)