$$S_K = \sum_{k=1}^{K} \dot{w}_k \tag{4.9}$$

信頼水準 99% の VaR は，1% を下回る最大の S_K と 1% を上回る最小の S_K を見つけ，それぞれ $S_{K(\mathrm{down})}$, $S_{K(\mathrm{up})}$ とすると，下式のようになる．

$$\mathrm{VaR}_{99} = \frac{(\dot{x}_{K(\mathrm{up})} - \dot{x}_{K(\mathrm{down})})\cdot(0.01 - S_{K(\mathrm{down})})}{S_{K(\mathrm{up})} - S_{K(\mathrm{down})}} \tag{4.10}$$

4.3 保有期間の変換

ポートフォリオのリスクを計測するとき，まず投資の目標となる期間を定める必要がある．分析目的にあった期間が選択されるが，例えば BIS 規制では 10 日間と決められている．これを保有期間という．しかし，リスクファクターの価格の多くは日次で得られるデータであり，これをもとに日次収益率が算出される．そのため 1 日の単位の収益率をもとに，10 日収益率を算出する必要がある．

この方法について 3 つのバリエーションが提案されている．**Box-Car 法, Moving-Window 法，ルート *t* 倍法**である．いずれの方法も一長一短あり，BIS 規制においても各銀行の開発者が選択していいことになっている．そのため，モデル開発者はそれぞれの手法の長短を認識し，目的やデータの質に合った無理のない方法を選択すべきである．以下に各手法の長所短所を中心に説明する．なお，保有期間がデータのサイクルと同じである場合はこのような処理は必要がない．

4.3.1 Box-Car 法

Box-Car 法は 10 日収益率を計算するのに，10 日分のデータを利用して計算するという，きわめて自然な考え方である．そのため，データに対する仮定は必要なく，どのようなデータ，目的に対しても利用可能である．

$$X_\tau^{(10)} = \sum_{t=10\tau-9}^{10\tau} x_t \tag{4.11}$$

$X_\tau^{(10)}$：リスクファクターの 10 日収益率

x_t：リスクファクターの日次収益率

問題点は分析精度を上げるために相当のデータ量を必要とすることである．図

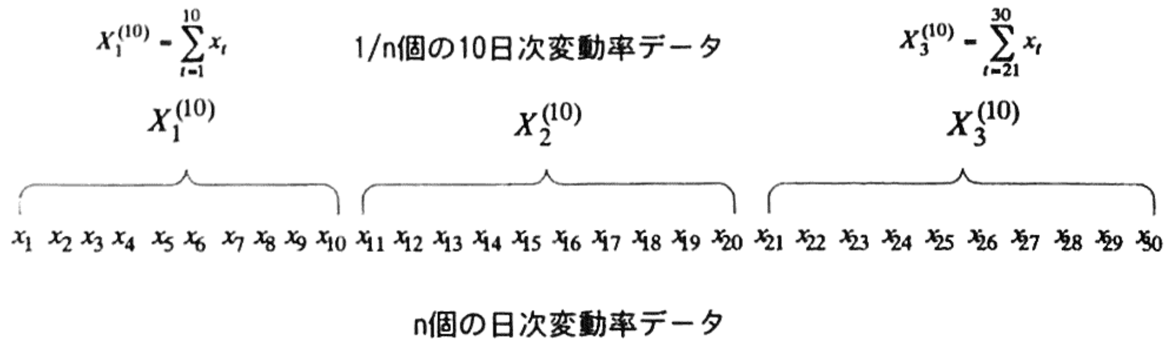

図 4.3　Box-Car 法による保有期間の変換

4.3 に 30 日間のデータで 10 日収益率を計算する方法を概念的に示したが，Box-Car 法だと 30 日間データより 3 つの 10 日収益率しか得られない．BIS 規制ではデータの採取期間（公式には観測期間）は 250 日以上 750 日以下であるため，250 日のデータでは 25 個の 10 日収益率データしか得られない．25 個のデータで，1% の下方リスク（VaR）を計算することが適当でないことは，感覚的に理解していただけると思う．

4.3.2　Moving-Window 法

Moving-Window 法は Box-Car 法のデータの減少を解決するため考案された方法で，収益率を観測するのに，保有期間のオーバーラップを認める方法である．そのため **Overlapping 法**と呼ばれることもある．この方法ならば，図 4.4 のように 30 日間のデータなら 21 個の，250 日データならば 241 個の 10 日間収益率データが得られる．

$$X_\tau^{(10)} = \sum_{t=\tau}^{\tau+9} x_t \tag{4.12}$$

しかし，隣り合う 2 つの 10 日間収益率データは，10 日のうち 9 日までが同じ日のデータをもとに計算されている．そのためデータ間の相関が高く，独立であるとはいえない．VaR を収益率分布から計算するときには，分布の各要素がそれぞれ独立であるのを前提としている．そのため，Moving-Window 法は統計的学的に妥当な計算プロセスだとはいえない．それでもこの方法が広く使われている背景には，有効なデータの少なさがある．つまり，古いデータは市場構造に違いがあるため現在の VaR の計測にはそぐわないが，Box-Car 法ではデータの

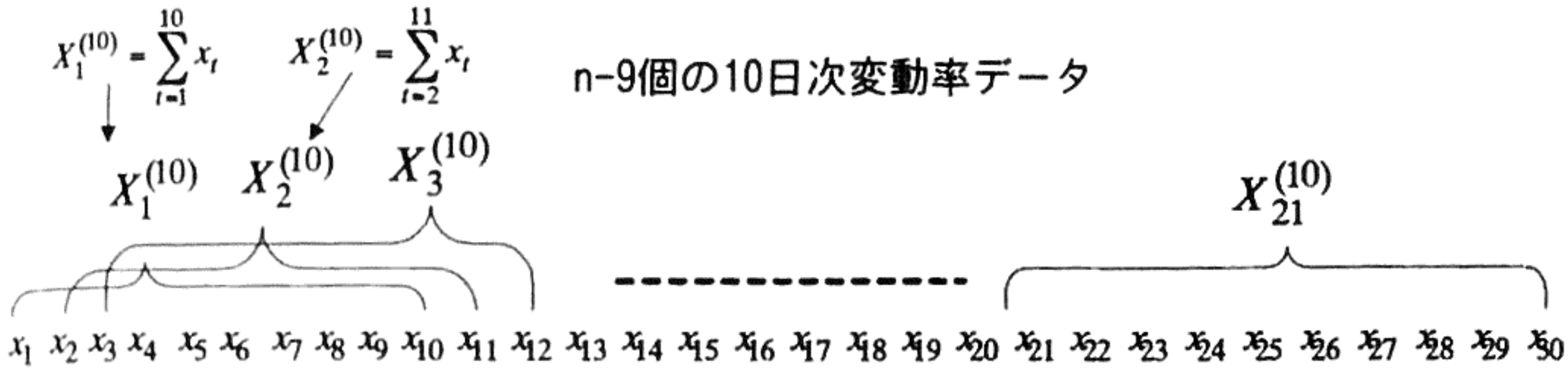

図 4.4 Moving-Window 法による保有期間の変換

少なさから正確に計算されないという，ジレンマの結果の選択であろう．ただ，一見データ数が多くなったように思えるこの方法も統計的にはデータの持つ情報量が増えるわけではなく，データの少なさが解決されることはない．この方法の問題点はむしろ，モデルの開発者がこの方法の欠点を理解しておらず，実際はモデルが不安定なのに，いかにも十分なデータによりモデルを開発したと思いこみがちなところにある．

4.3.3 ルート t 倍法

ルート t 倍法も Moving-Window 法と同じく，Box-Car 法の取得されるデータの少なさを解決するため考案された方法である．この方法は，あらかじめ日次データのデータセットで推定された日次のボラティリティに基づき，10 日間ボラティリティを計算する方法である．この方法は，日次データに系列相関がない，言い換えれば日次の収益率が時系列的に独立であるという前提を用いている．この前提が成り立つとき，10 日間のボラティリティ（分散）は，日次データのボラティリティのルート 10 倍となるという性質を利用したものである．一般的に t 日間ボラティリティは日次の $\sqrt{t}$ 倍なので，ルート t 倍法と呼ばれている．

時系列的に独立であると，分散が $\sqrt{t}$ 倍で表されるということは，以下のように導くことができる．

今，対数価格 $\ln P_t$ が，次のような過程（ランダムウォーク）で変動すると仮定する．

$$\ln P_t = \delta_t^{(1)} + \ln P_{t-1} + \varepsilon_t^{(1)} \tag{4.13}$$

ここで，$\delta_t^{(1)}$ は非ランダムドリフトパラメータで，分散は 0 である．また添え字の (1) は，1 日の変動という意味である．$\varepsilon_t^{(1)}$ は平均 0，分散 σ_1^2 の独立および同一の正規分布の確率変数と仮定する．すると，

$$\ln(P_t/P_{t-1}) = \delta_t^{(1)} + \varepsilon_t^{(1)} \tag{4.14}$$

となり，$\varepsilon_t^{(1)}$ の分布関数がそのまま収益率分布を表すモデルとなる．

上式で再帰的代入を行うと次の関係式を得る．

$$\ln(P_{10}/P_0) = \sum_{t=1}^{10} \delta_t^{(1)} + \sum_{t=1}^{10} \varepsilon_t^{(1)} \tag{4.15}$$

この 10 日間の対数収益率の期待値は，$E[\varepsilon_t^{(1)}] = 0$ より次式で与えられる．

$$E[\ln(P_{10}/P_0)] = \sum_{t=1}^{10} \delta_t^{(1)} \tag{4.16}$$

この式が与えられれば，10 日間の対数収益率の分散は，次式で与えられる．

$$V[\ln(P_{10}/P_0)] = E\left[\left(\sum_{i=1}^{10} \varepsilon_t^{(1)}\right)^2\right] = 10\sigma_1^2 \tag{4.17}$$

したがって，収益率の標準偏差に関して次の関係式が成立する．

$$\sigma^{(10)} = \sqrt{10}\sigma_1 \tag{4.18}$$

この関係が成立するのは，10 日間のランダムウォークが平均 0，分散が 10 日間毎日同じ σ_1^2 の分布に従い，かつ独立なときに限られる．

モデル誤差項に有意な系列相関が存在する場合は，このルールが適用できない．そのため収益率データセットに，ある時系列モデルを用い，相関を取り除いたものをランダムウォークとしてとらえる必要がある．しかし，現在リスク管理モデルとして実務的に用いられているヒストリカル法やデルタ法では，収益率データセットをそのままランダムウォークとしてとらえている．

ルート t 倍ルールの信頼性を判断するためには，収益率が系列相関があるかどうかを調べればよい．収益率の観測値 $x_{i,t}$ について，ラグ数 l に対し，自己相関係数 ρ（l 期間ずつ離れた観測値の自己相関の値）を求める．

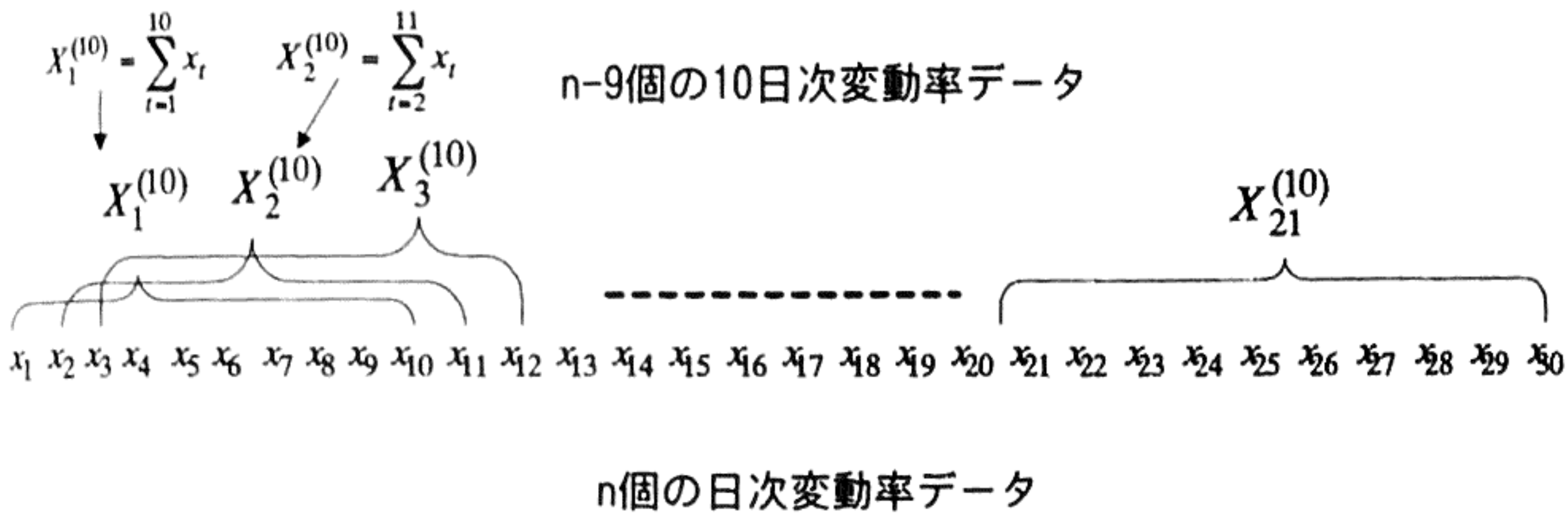

図4.4　Moving-Window 法による保有期間の変換

少なさから正確に計算されないという，ジレンマの結果の選択であろう．ただ，一見データ数が多くなったように思えるこの方法も統計的にはデータの持つ情報量が増えるわけではなく，データの少なさが解決されることはない．この方法の問題点はむしろ，モデルの開発者がこの方法の欠点を理解しておらず，実際はモデルが不安定なのに，いかにも十分なデータによりモデルを開発したと思いこみがちなところにある．

4.3.3　ルート t 倍法

ルート t 倍法も Moving-Window 法と同じく，Box-Car 法の取得されるデータの少なさを解決するため考案された方法である．この方法は，あらかじめ日次データのデータセットで推定された日次のボラティリティに基づき，10日間ボラティリティを計算する方法である．この方法は，日次データに系列相関がない，言い換えれば日次の収益率が時系列的に独立であるという前提を用いている．この前提が成り立つとき，10日間のボラティリティ（分散）は，日次データのボラティリティのルート10倍となるという性質を利用したものである．一般的に t 日間ボラティリティは日次の $\sqrt{t}$ 倍なので，ルート t 倍法と呼ばれている．

時系列的に独立であると，分散が $\sqrt{t}$ 倍で表されるということは，以下のように導くことができる．

今，対数価格 $\ln P_t$ が，次のような過程（ランダムウォーク）で変動すると仮定する．

$$\ln P_t = \delta_t^{(1)} + \ln P_{t-1} + \varepsilon_t^{(1)} \tag{4.13}$$

ここで，$\delta_t^{(1)}$ は非ランダムドリフトパラメータで，分散は0である．また添え字の(1)は，1日の変動という意味である．$\varepsilon_t^{(1)}$ は平均0，分散 σ_1^2 の独立および同一の正規分布の確率変数と仮定する．すると，

$$\ln(P_t/P_{t-1}) = \delta_t^{(1)} + \varepsilon_t^{(1)} \tag{4.14}$$

となり，$\varepsilon_t^{(1)}$ の分布関数がそのまま収益率分布を表すモデルとなる．

上式で再帰的代入を行うと次の関係式を得る．

$$\ln(P_{10}/P_0) = \sum_{t=1}^{10} \delta_t^{(1)} + \sum_{t=1}^{10} \varepsilon_t^{(1)} \tag{4.15}$$

この10日間の対数収益率の期待値は，$E[\varepsilon_t^{(1)}] = 0$ より次式で与えられる．

$$E[\ln(P_{10}/P_0)] = \sum_{t=1}^{10} \delta_t^{(1)} \tag{4.16}$$

この式が与えられれば，10日間の対数収益率の分散は，次式で与えられる．

$$V[\ln(P_{10}/P_0)] = E\left[\left(\sum_{i=1}^{10} \varepsilon_t^{(1)}\right)^2\right] = 10\sigma_1^2 \tag{4.17}$$

したがって，収益率の標準偏差に関して次の関係式が成立する．

$$\sigma^{(10)} = \sqrt{10}\sigma_1 \tag{4.18}$$

この関係が成立するのは，10日間のランダムウォークが平均0，分散が10日間毎日同じ σ_1^2 の分布に従い，かつ独立なときに限られる．

モデル誤差項に有意な系列相関が存在する場合は，このルールが適用できない．そのため収益率データセットに，ある時系列モデルを用い，相関を取り除いたものをランダムウォークとしてとらえる必要がある．しかし，現在リスク管理モデルとして実務的に用いられているヒストリカル法やデルタ法では，収益率データセットをそのままランダムウォークとしてとらえている．

ルート t 倍ルールの信頼性を判断するためには，収益率が系列相関があるかどうかを調べればよい．収益率の観測値 $x_{i,t}$ について，ラグ数 l に対し，自己相関係数 ρ（l 期間ずつ離れた観測値の自己相関の値）を求める．

$$\rho_l = \frac{\sum_{t=l+1}^{T}(x_t - \bar{x})(x_{t-j} - \bar{x})/[T-(l+1)]}{\sum_{t=1}^{T}(x_t - \bar{x})^2/(T-1)} \tag{4.19}$$

$$\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t \tag{4.20}$$

ここで，T は系列 x_t の観測値の数を表す．

複数の次数の自己相関を同時に検定する手法がある．その1つが **Box-Ljung 検定**である（Ljung and Box (1978)）．統計量は次のようになる．

$$10(10+2)\sum_{l}\frac{\rho_l^2}{10-1} \tag{4.21}$$

RiskMetrics では，Box-Ljung 検定の結果，いくつかの資産クラスでは系列相関があることが判明している．したがってそういった資産クラスで，ルート t 倍ルールや Moving-Window 法を使うことには問題がある．しかしながら，自己相関の影響は強くなく，また，自己相関をモデル化するのは非常に難しいという理由から，系列相関を考慮に入れた調整は行っていない．実務的にも，収益率に系列相関がないものと仮定してモデル作成が行われている．

最後に表 4.1，4.2 に 3 つの方法によって算出した TOPIX のボラティリティと 99%VaR を示した．方法の違いによる算出結果に時として大きな差異があることがわかる．

表 4.1 保有期間の処理法による標準偏差（%）の違い
TOPIX 日次データ（1988～1999）より作成
（観測期間 250 日，保有期間 10 日，ウエイトなし）．

観測期間	Box-Car 法	Moving-Window 法	ルート t 倍法
88/01/04～88/11/29	2.57	2.50	2.51
88/11/30～89/11/29	2.40	1.97	1.86
89/11/30～90/12/03	4.32	5.46	5.36
90/12/04～91/12/09	3.53	3.77	3.58
91/12/10～92/12/14	6.98	5.88	5.00
92/12/15～93/12/20	3.64	4.05	3.74
93/12/21～94/12/26	2.87	2.55	2.86
94/12/27～95/12/26	4.09	3.71	3.68
95/12/27～97/01/05	2.65	2.48	2.38
97/01/06～98/01/11	3.75	3.49	4.40
98/01/12～99/01/13	4.51	3.93	4.39
（平均値）	3.75	3.62	3.61

表 4.2　保有期間の処理法による VaR の違い
10000 円の保有に対する 99%VaR を示す
(単位円，デルタ法，ウエイトなし，観測期間等は表 4.1 と同じ).

観測期間	Box-Car 法	Moving-Window 法	ルート t 倍法
88/01/04〜88/11/29	572.5	556.7	559.7
88/11/30〜89/11/29	534.4	439.7	414.0
89/11/30〜90/12/03	963.8	1217.0	1196.0
90/12/04〜91/12/09	787.4	840.4	797.3
91/12/10〜92/12/14	1556.1	1311.5	1116.1
92/12/15〜93/12/20	810.9	903.2	833.0
93/12/21〜94/12/26	640.3	569.3	637.3
94/12/27〜95/12/26	912.1	828.0	821.7
95/12/27〜97/01/05	591.6	553.5	530.5
97/01/06〜98/01/11	836.7	777.6	980.4
98/01/12〜99/01/13	1004.6	877.2	978.3
(平均値)	837.31	806.74	805.86

BOX-9　Box-Car 法とヒストリカル法の組合せ

Box-Car 法の場合，分析に使用するデータ数が少ないので 99% 点（最大損失収益率）を代表する値が存在しないモデルが発生する．以下に，それぞれの使用データ数ごとに 99% 点を推定する手法を述べる．

① データ数=250 の場合，分析に使用するデータ数は 25 であり，一番小さい収益率 $\dot{x}_1$ の階級値は $0+1/(25\times 2)=2\%$ である．そのため 99%VaR を求めることができない．やむを得ないときは一番小さい収益率で 1% 点で近似する．

② データ数=500 の場合，分析に使用するデータ数は 50 であり，一番小さい収益率 $\dot{x}_1$ の階級値は $0+1/(50\times 2)=1\%$ である．したがって，一番小さい収益率で 1% 点を推定する．

③ データ数=750 の場合，分析に使用するデータ数は 75 であり，一番小さい収益率 $\dot{x}_1$ の階級値は $0+1/(75\times 2)=0.67\%$，また 2 番目に小さい収益率 $\dot{x}_2$ の階級値は $1/75+1/(75\times 2)=2\%$ である．したがって $\dot{x}_1, \dot{x}_2$ の収益率を線形補間した値で 1% 点を推定する．具体的には次の式で求められる値を用いる．

$$\frac{\dot{x}_1(2-1)+\dot{x}_2(1-0.67)}{(1-0.67)+(2-1)}=0.75\dot{x}_1+0.25\dot{x}_2$$

しかし，一番小さい収益率 $\dot{x}_1$ は非常にセンシティブな数値である．例えば，デー

タの中に大暴落などの異常値が1つ存在すると，x_1 が反応し分析結果 VaR が大きく変動する．そのため，この方法は極端な値を出す傾向にあり，分析結果を投資戦略の意思決定として使うことが難しい．結論としては，ヒストリカル法を採用した場合，保有期間の処理はルート t 倍法もしくは Moving-Window 法で行うべきであろう．

4.4 欠損データの処理とデータ取得タイミング

4.4.1 欠損データの処理

統計学においては，データの補間に関する様々な方法が存在する．しかし，VaR の算出においては，データの補間を厳密に行ったからといって，さほど全体の精度向上に貢献しないとの報告がある．そこで，一般には直近のデータによる補間などの簡便な方法が採用されることが多い．これ以外にも，少し厳密な方法として線形補間，最も厳密な方法として EM アルゴリズムがある．逆に最も単純に，ヒストリカル法やデルタ法においては，欠損データがある日のデータは補間をせずにすべて削除する方法も考えられる．以下に，代表的な欠損値処理法について述べる．

a. 直前のデータを使用する

他の補間方法はデータ欠損日以降に次のデータが入手されて計算が可能となるが，この方法はそれ以降のデータを待たずに，当日にデータを補間することが可能であるという利点がある．この方法は，休日等の理由以外では欠損データが生じる可能性はきわめて少ないことや，実際に把握していない市場の動きについて，不完全な形で修正データを作成する可能性がないこと（恣意性の排除），ま

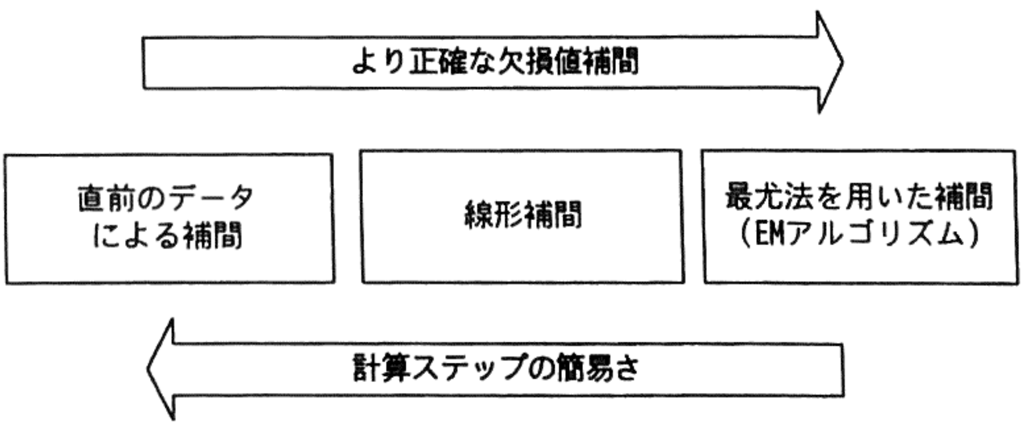

図 4.5 欠損値補間のバリエーション

た補間による分散へのインパクトが大きくないなどの判断から，合理的である．さらに，計算ステップがなく簡便であることも大きな利点である．

なお，一般的なシステムでは，どのデータが前日補間されたかをチェックするため「前日補間リスト」が出力される．何らかのアクシデントにより，正しいデータが存在するにもかかわらず，必要データが欠落している場合は，手入力による修正を行うことにより，データ欠損の影響を少なくできる．

b. 前後のデータで線形補間する

欠損当日は前日のデータをそのまま適用するが，その後の営業日にデータを取得できた時点で欠損の前営業日と取得できた時点との間で線形補間を行う．

$$\hat{x}_t = \frac{1}{2}x_{t-1} + \frac{1}{2}x_{t+1} \qquad \text{（データ欠損が 1 期間）} \tag{4.22}$$

$$\hat{x}_{T+t} = \frac{N-t}{N}x_T + \frac{t}{N}x_{T+N} \qquad \text{（データ欠損が } N-1 \text{ 期間）} \tag{4.23}$$

c. EM アルゴリズム等で推定する

データが多次元正規分布に従っていることを前提とし，最尤推定法を用いて欠損データを求める方法の 1 つが，EM アルゴリズム（expected maximum algorithm）である．

もともと EM アルゴリズムは，不完全なデータセットからモデルパラメータの最尤推定を行う手法で，その過程で欠損データの補間が行われる．この方法は近年の情報幾何学の発展により，その理論的体系や計算方法，改良アルゴリズムなどが構築されている．そのため，いくつかのバリエーションがあるが，その基本的な概念は以下のとおりである．

EM アルゴリズムは E-step（expectation step）と M-step（maximization step）の 2 つの部分からなり，これらを交互に繰り返してパラメータを更新することによって，最尤推定量あるいは尤度関数の極大点を得る．

① 初期パラメータを与えたモデルを作成する．

② モデルを用いて欠損値の推定を行う．（E-step）

③ 欠損値を補間されたデータセットを用いて，モデルのパラメータを最尤推定し，更新する．（M-step）

④ 上記の ②③ のステップを繰り返すことにより，尤度関数の極大化を行う．

実際の EM アルゴリズムの運用については，甘利（1992）などを参照されたい．

なお，EM アルゴリズムをはじめとする，最尤データ補間法は，直近値補間，線形補間に比較して合理性が高いと判断されている．

異常値判定： 補間を行った場合，その補間が適当であったかどうか判定することが必要となる．これは一般に用いられている異常値の判定と同様の方法でもよいが，補間という作業に特徴的な傾向を念頭に置き，チェックした方がよい．時系列データの補間に異常データが含まれると，

① リターン分布の尖度が異常に大きくなる．

② 相関係数が大きく異なったものとなる．

③ 金利データについてはイールドカーブの形状変化が通常のものと異なる．

などの特徴がみられる．

4.4.2 データ取得のタイミング

データの取得のタイミングをどのようにするかという問題については，基本的に以下の 2 つの方法がある（図 4.6）．

① 各市場のデータを同一のタイミングで取得する．

② 各市場の終値データを取得する．

異なるマーケット間の相関をできるだけ正確に推計するためには，どの時点でマーケットデータを採取するかが重要である．とりわけ日次でデータを取得し，

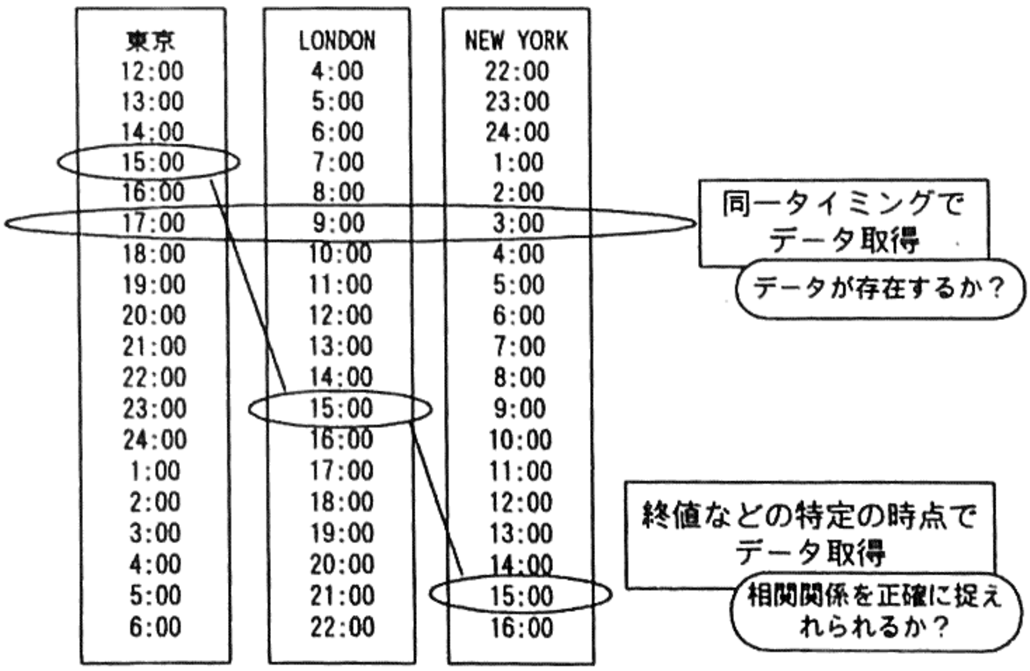

図 4.6 データ取得タイミングのバリエーション

相関係数を推計する場合はそのタイミング（時刻）を考慮する必要がある．

データを取得するタイミングを決めるにあたっては，リスク管理上最も重要な日本市場を中心として，極力多くのマーケットが開いている同時点でデータを取得することが望ましい．それは，金融マーケットではあるニュースが各国市場にほぼ同じタイミングで伝播し，瞬時にそのニュースを織り込んでしまうと考えられるからである．それゆえ，できるだけ多くのマーケットが開いている時刻でデータを採取することにより，マーケット相互の相関係数の精度をより高めることができる．逆にマーケット間でデータ採取時間が異なれば，その間に発生したニュースを織り込んだ市場と織り込まない市場がでてくるので，マーケットの相関をみるには好ましくない．一般的に相関が現実より低く計算される．

しかし，すべてのマーケットが同時に開かれていないことや，たとえ開かれていたとしても同時採取が困難なこともある．例えば東京市場における欧州債券のように，商品・種類によっては市場の厚みが十分でなく，適切な市場動向を反映したデータ採取を行うことができない可能性がある．また，市場の寄りつきの価格を使用する場合も，同様に市場の厚みが不十分であることから，指標性に問題がある．これらを比較すると，データによっては各市場の清算値あるいは終値を使用することは現実的な方法といえる．しかしながら，この方法はあくまでも簡便法であるため，採取したデータに明らかな偏りが観測される場合には，適宜その対応策を練り，データの客観性・指標性を維持していく必要がある．

章末問題

問題 4.1 ルート t 倍法は日次収益率が独立であるという仮定で成り立っていた．もし，独立が保証されない場合，99%VaR は最大どのくらいになるか．ただし日次収益率の標準偏差を σ としデルタ法を前提とせよ．

問題 4.2 減衰係数が 0.7 の指数加重移動平均法を行ったとする．観測期間を 10 日（$T-0$ 日目～$T-9$ 日目）にした場合の各データに対するウエイト $w_0, w_1, \cdots, w_9$ を求めよ．そのとき 11 日以前のウエイトの累積和 Ω_{T_L} を求めよ．また，残分ウエイト累積和 Ω_{T_L} が大きいと判断したときの修正ウエイト $w_0', w_1', \cdots, w_9'$ を計算せよ．