

サンプル操作をマスターする

株式会社ライトストーン

これまでも EViews での利用サンプルの変更方法については折に触れ解説してきました。非常に重要な操作ですので、今回は EViews 上でのサンプル操作について、応用的な点も含めまとめます。

1 基本的な使い方について

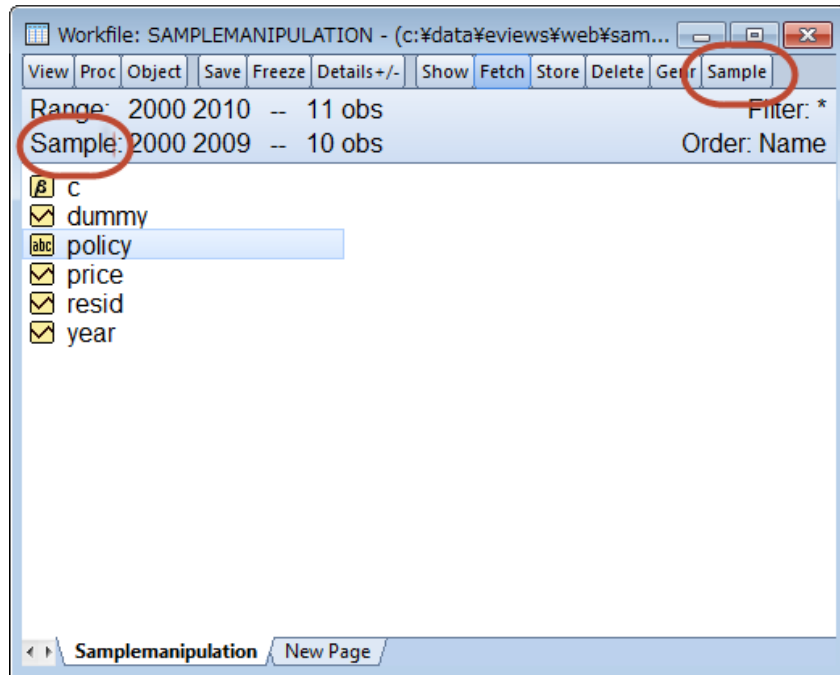
実証分析を行う際、データの部分集合に対して分析を行い、結果を比較したいことがよくあります。例えば男女混合の調査から、男性だけ・女性だけを取り出して傾向をグラフ化したり、時系列データをバブル前とバブル後に分割して分析を行いたいような場合です。このような場合、利用する統計ソフトウェアによっては、不必要なデータを実際に（一時的に）削除する必要があります。これには誤操作によりデータが永久に失われてしまう危険性や、データファイルが複数になってしまうことによる利便性の低下が伴います。

EViews は、観測値 (Observation) とサンプル (Sample) を明確に区別することに一つの特徴があります。EViews における“観測値”はデータの全体を表し、原則的に変更できません。その代わりに、分析に利用するサンプルを条件指定することで、観測値の中の分析の対象を狭めることが可能です。この特徴により、データの部分集合を取ってそれぞれに対して分析を行いたいような場合に、別途用意したデータを読み込み直したりする必要がありません。また、サンプルは簡単に元に戻せますので、元々のデータが失われてしまう危険も基本的にありません。

今回はこのサンプル操作についてまとめます。データとして、弊社で作成した SampleManipulation.wf を用います。まず基本的な操作から確認していきます。サンプルのセットはメニュー操作とコマンド操作のどちらでも行うことができますが、両者には入力内容で微妙な違いがありますので、その点も含め解説します。全体に先立ち、コマンド `smpl @all` で、サンプルを元に戻せることに触れておきたいと思います。効果を分かりやすくするために、各操作を試す前に `smpl @all` を毎回実行してください。このコマンドは実際の分析においても非常によく使いますので、覚えておいて損はありません。

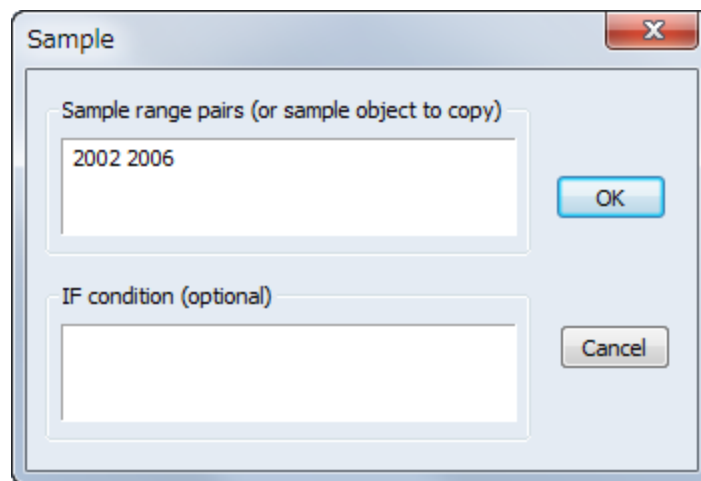
1.1 メニュー操作について

まずメニュー操作を行いたい場合について解説します。以下のスクリーンショットの場所のどちらかをクリックします。



もしくは、メニューより Proc > Set Sample と操作します（どのように操作しても、表示されるダイアログは同じです）。

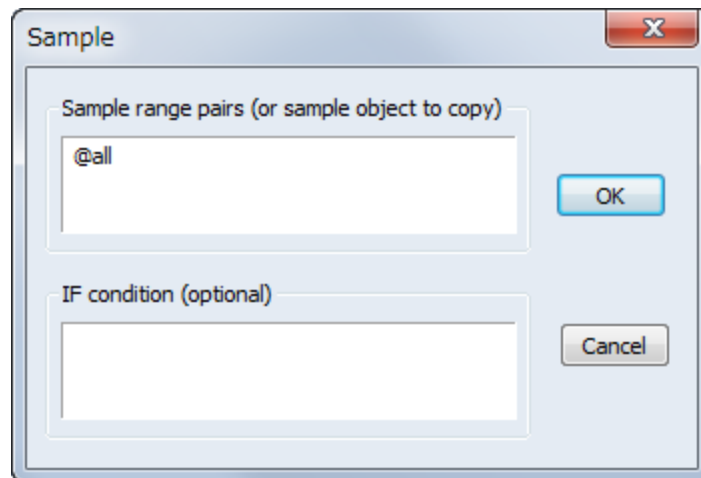
ポップアップしたダイアログに以下のように入力して OK してください。



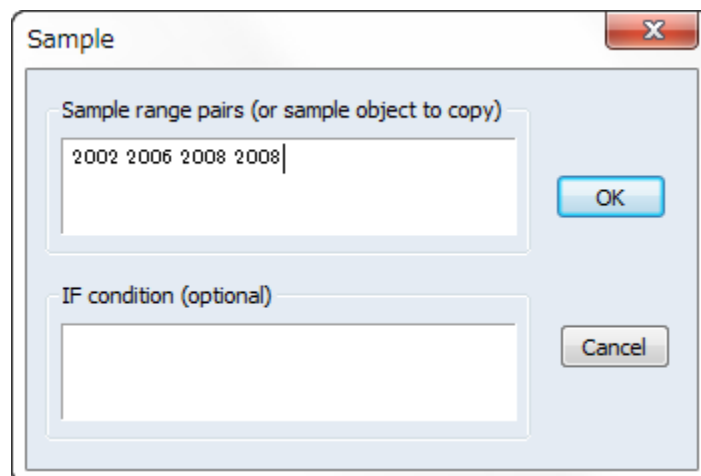
すると、以下のようにサンプル表記が置き換わります。

Range: 2000 2010 -- 11 obs
 Sample: 2002 2006 -- 5 obs

この状態では、2002 年から 2006 年までのデータのみを用いて、統計分析が行われることになります。これを元に戻したい場合は、再度上記の操作を行い、“2000 2010”と入力します。もしくは、次のように@all と入力します。



2002年から2006年に加え、2008年を単独で使いたいような場合、以下のように入力します。



ポイントは以下の通りです。

- 期間を複数設定したい場合、それぞれの期間毎のペアを追加していきます
- ある1年だけを追加したい場合も、単に一回入力すればよいわけではなく、必ず“2008 2008”などとペアで入力する必要があります

1.2 コマンド操作について

上記の操作をコマンドベースで行いたい場合、以下のように入力します。

```
smp1 2002 2006
```

あるいは、この場合、シリーズ `year` で時間が定義されていますので、`if` 条件を使ってサンプルを限定することも考えられます。試しに次のコマンドを連続で実行してみてください。

```
smp1 @all if year >= 2002
```

```
smpl @all if year <= 2006
```

確認してみると、最後に実行した `smpl @all if year <= 2006` のみが有効になっており、2002 年以降という最初の条件が無効になっていることが分かります。これは、`@all` に対して、改めて 2006 年以降という条件が設定されているためです。したがって、たとえば以下のように入力すれば、利用サンプルを 2002 年から 2006 年までに設定することができます。

```
smpl 2002 2010 if year <= 2006
```

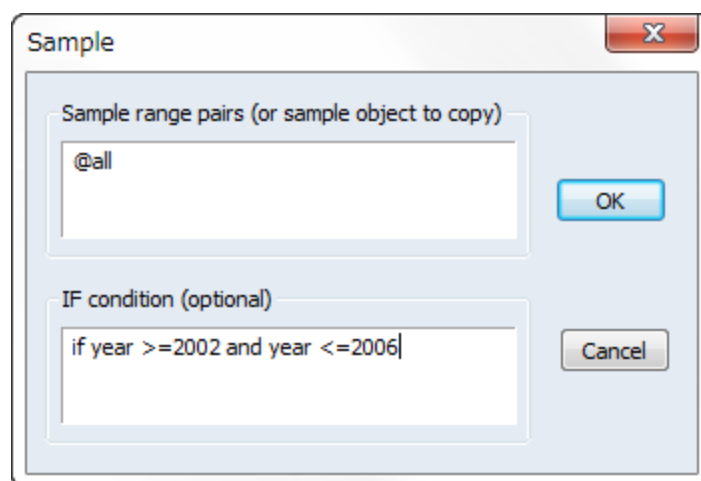
ただしこの場合、`if` 条件にすべての条件を記述したほうが、よりスマートでしょう。その場合は以下のように入力します。

```
smpl @all if year >= 2002 and year <= 2006
```

ここでの重要な点は、上記をまとめて、`smpl @all if 2002 <= year <= 2006` などとすることはできないということです。仮にこのように記述した場合、`if` 以降が評価されず、`smpl @all` という表記と同じ処理が実行されることとなりますので、お気を付けください。

1.3 メニュー操作とコマンド操作の違い

基本的には、メニュー操作とコマンド操作で、操作にそれほど違いはありません（メニュー操作でも、最終的には期間などをキーボード入力する必要があるからです）。メニュー操作とコマンド操作の大きな違いは、コマンド操作の場合、`smpl` や `if` を入れてはいけないということです。試しに、サンプルダイアログの Sample Range Pairs 入力フィールドに、`smpl 2002 2006` と入力してみてください。また、以下のように入力してみてください。



この状態で OK ボタンを押すと、エラーが生じます。このようにメニュー操作の場合、`smpl` や `if` を省略した記述を行わなければなりません。

2 応用的な使い方について

文字列による条件付け アルファシリーズオブジェクト policy には、A か B どちらかの文字列が格納されています。文字列でサンプルを分割したい場合は以下のようにします。

```
smpl @all if policy = "A"
```

違いは、文字列の場合、ダブルクォーテーションで入力値を囲う必要があるということです。

数式表現 条件に数式表現を利用することも可能です。たとえば、Price が平均以上のデータだけを使って分析したい場合は以下のようにします。

```
smpl @all if price >= @mean(price)
```

もちろん、平均以外の関数を使うことも可能です。EViews で利用可能な関数については、EViews Help Topics から、Operator and Function Listingなどを参考にしてください。

便利な表現 財務データなどの定期的に更新がなされるデータを用いて分析を行う時、分析の度に新しい期間を指定し直すのは不便である場合があります。

たとえば観測値の最初と最後の2年を取り除いて分析を行いたい場合を考えます。このような場合、以下のように記述します。

```
smpl @first+2 @last-2
```

このように@first, @last といった表現を使うことによって、それぞれ観測値の最初と最後を指定することができます。さらに、プラスとマイナスによって調整しています。もちろん、現在使っているワークファイルでは単に smpl 2002 2008 と入力しても同じ出力が得られますが、観測値が年次更新され、2011年のデータが追加されたような場合を考えると、両者の出力が異なることが分かります。同じ分析を定期的に繰り返す場合に便利です。

複雑な if if 条件を and や or を複数組み合わせることで記述したい場合、条件が考慮される順序によって、出力が意図せず変わってしまう場合があります。これを防ぐため、条件が複雑な場合は () を用います。

```
smpl @all if (year >= 2002 and year <= 2006) or dummy = 1  
smpl @all if year >= 2002 and (year <= 2006 or dummy = 1)
```

上の表現では、「year の値が 2002 以上 2006 以下」もしくは「dummy の値が 1」という条件になるのに対し、下の表現では「year の値が 2002 以上」かつ「year の値が 2006 以下もしくは dummy の値が 1」という条件になります。実際に出力も異なります（上の表現では 2000 年が入るのに対し、下の表現では入りません）。括弧の中身から先に解釈されますので、複雑な if 条件を作りたい場合は、順序が曖昧にならないように適宜括弧を入力してください。

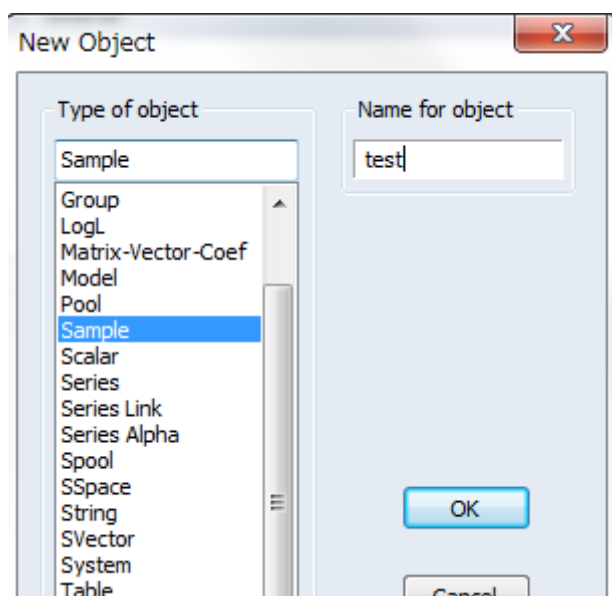
あえて観測値を変更したい 本書の最初に、EViews ではサンプルを変更し、観測値は変更しないという原則を書きました。あえてなんらかの理由で観測値そのものを変更したい場合、pagestruct コマンド (ワークファイルウィンドウの Proc>Structure/Resize Current Page) や pagecontract コマンド (ワークファイルウィンドウの Proc>Contract Current Page) を利用します。事前にバックアップを取っておくことを強くお勧めします。

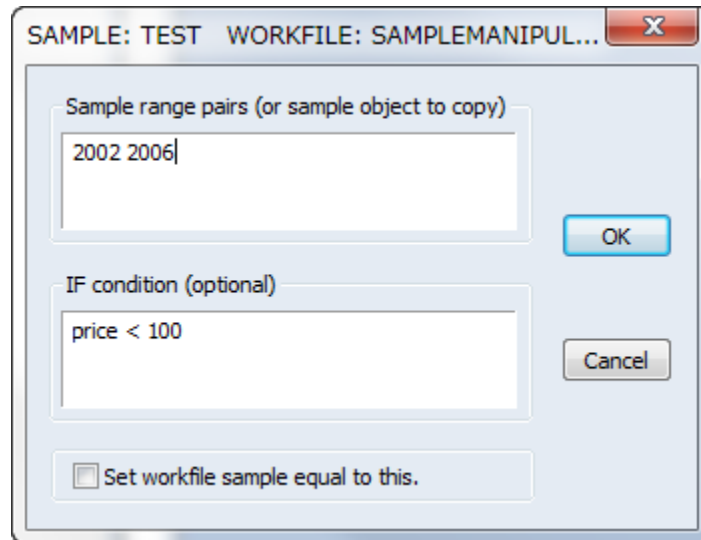
現在のサンプルのみを外部ファイルにエクスポートしたい データを外部ファイルにエクスポートする際、EViews はデフォルトで現在選択中のサンプルではなくデータの全体を保存します。外部アプリケーションとの連携のために現在選択中のサンプルのみをエクスポートしたい場合、以下のようにします。

1. pagesave コマンド (あるいは wfsave 等の類似のコマンド) を実行します。
2. “ファイルの種類” を、EViews Workfile 以外 (例えば、Excel File) に変更します。
3. Sample to write タブを開き、@all を本書の内容を元書き換えます (例えば 2002 2009 とすれば 2002 年から 2009 年までのデータのみが保存されます)。
4. エクスポートする変数を指定した後、OK を押します。

3 サンプルオブジェクトについて

ここまで何度もサンプルの変更操作を行ってきましたが、特定のサンプル操作を頻繁に行う必要がある場合、同じ条件を毎回入力するのは面倒です。このような場合、サンプルオブジェクトを作ることで、煩雑な条件を入力するのは一度で済みます。サンプルオブジェクトは、その名の通りオブジェクトの一種ですので、右クリックから New Object を選択した後、以下のようにして作成します。





ただし、これだけではサンプルに変更は起こりません。設定したサンプルオブジェクトを適用するためには、上の画面で“Set workfile sample equal to this”にチェックを入れます（サンプルオブジェクトを作成後は、Procメニューよりアクセスできます）。

あるいは、コマンド操作の際に以下のように用いることも可能です。

```
smpl test
```

コマンドウィンドウに smpl コマンドを入力した場合、EViews を開いている間は過去に実行したコマンドがウィンドウに表示されていますが、EViews を閉じるとそれらの情報は消えてしまいます。一方で、一度サンプルオブジェクトを作って保存しまえば、ワークファイルを閉じてもオブジェクトとして条件が残り続けます。サンプルオブジェクトはこの点で便利です。

サンプルの設定は日常的な操作ですが、if 条件の記述を間違った場合のように、エラーが起きなくても条件が意図した通りに設定されていない場合があります。まずは正しくサンプルが正しく変更されているかどうかを、目視で確認することを忘れないようにしてください。