

平成 28 年 7 月 29 日
 新旧データ接続検討WG

現時点における計測方法の候補

1. 新標本と旧標本の双方において、同一時点・同一事項を調査している場合、新旧双方の標本で同一時点の平均等の統計量を計算し、その差異をサンプル替えの影響とする。

<例>

- 毎月勤労統計では、標本交替の際、新旧標本の双方で1月の状態を調査
- 法人企業統計では、期首と期末のストックを調査していることから、標本交替に際しても、旧標本で期末の状態、新標本で期首の状態を把握

2. 新標本と旧標本の差の算式を変形し、要因分解の形式にしてサンプル変更要因と解釈可能な特定の項をサンプリング変更の影響とみなす方法

<例>

x : 新標本 y : 旧標本、 θ 、 ϕ : 各種パラメータ等

$$\begin{aligned} \text{統計量の変動} &= f(x, \theta) - f(y, \phi) \\ &= f(x, \theta) - f(y, \theta) + f(y, \theta) - f(y, \phi) \\ &= \underbrace{[f(x, \theta) - f(y, \theta)]}_{\text{標本変動要因項}} + \underbrace{[f(y, \theta) - f(y, \phi)]}_{\text{パラメータ変動要因項}} \end{aligned}$$

3. 標本交替により、売上高等の推定結果には標本交替前と標本交替後でレベルシフトが発生すると想定し、時系列解析ソフト（X-12 等）で、標本交替時におけるレベルシフトダミー要素を検出する方法

4. ①標本交替時に生じる変動（前期比、前年同月比）と、②標本交替を行わない時の変動（前期比、前年同月比）を比較すると、①の変動の方が大きいことが想定される。そのため、①の標本交替時の前期比・前年同期比と②の標本交替を行わない時の前期比・前年同期比の大きさを長期間のデータによって比較し、標本交替時の平均的な影響の大きさ（①から②を引いた値等）で検出する方法

5. シミュレーションを行いラフに影響を計算する方法

ア) N個の2時点 (t,t+1) の継続標本データを用意する。

$x(1,t)$	$x(1,t+1)$
$x(2,t)$	$x(2,t+1)$
$x(3,t)$	$x(3,t+1)$
$x(4,t)$	$x(4,t+1)$
\vdots	\vdots
$x(n,t)$	$x(n,t+1)$

イ) 「ア)」のデータから n 個の継続標本データをランダムに複数回取り出し、その都度平均を計算する

$x(i_1,t)$	$x(i_1,t+1)$
$x(i_2,t)$	$x(i_2,t+1)$
$x(i_3,t)$	$x(i_3,t+1)$
$x(i_4,t)$	$x(i_4,t+1)$
\vdots	\vdots
$x(i_n,t)$	$x(i_n,t+1)$

 $u(t)$ $u(t+1)$ ……抽出の都度、平均を計算

ウ) 「ア)」のデータから t 期、t+1 期それぞれから n 個のデータをランダムに複数回取り出し、その都度平均を計算する。

$x(j_1,t)$	$x(k_1,t+1)$
$x(j_2,t)$	$x(k_2,t+1)$
$x(j_3,t)$	$x(k_3,t+1)$
$x(j_4,t)$	$x(k_4,t+1)$
\vdots	\vdots
$x(j_n,t)$	$x(k_n,t+1)$

 $s(t)$ $s(t+1)$ ……抽出の都度、平均を計算

エ) 「イ)」で複数回計算される「 $u(t+1) - u(t)$ 」と「ウ)」で複数回計算される「 $s(t+1) - s(t)$ 」の標準偏差の差を調べる。