中国人民大学应用统计科学研究中心
中国人民大学统计学院

主　编　金勇进
副主编　田茂再　刘　畅

统计学评論
Statistical Review

Vol. 8

连续调查的差值估计法

贸易增加值率背后是企业增加值率
——针对中国加工贸易出口的讨论与测算分析

技术创新与非技术创新的协同发展：
基于珠三角企业转型升级路径的实证研究

新闻事件主题与股票收益率波动的联动分析

基于六西格玛原理的刹车材料生产流程改善
——以X公司为例

经济科学出版社
Economic Science Press

中国人民大学应用统计科学研究中心

中国人民大学统计学院

统计学评论

Statistical Review

Vol. 8

主　编　金勇进

副主编　田茂再　刘　畅

经济科学出版社

图书在版编目（CIP）数据

统计学评论．8/金勇进主编．—北京：经济科学出版社，2014．3
ISBN 978－7－5141－4434－5

Ⅰ．①统…　Ⅱ．①金…　Ⅲ．①统计学－文集　Ⅳ．①C8－53

中国版本图书馆 CIP 数据核字（2014）第 047603 号

责任编辑：刘怡斐
责任校对：王苗苗
责任印制：邱　天

统计学评论
Statistical Review
Vol. 8
主　编　金勇进
副主编　田茂再　刘　畅
经济科学出版社出版、发行　新华书店经销
社址：北京市海淀区阜成路甲 28 号　邮编：100142
总编部电话：010－88191217　发行部电话：010－88191522
网址：www. esp. com. cn
电子邮件：esp@ esp. com. cn
天猫网店：经济科学出版社旗舰店
网址：http：//jjkxcbs. tmall. com
北京万友印刷有限公司印装
787×1092　16 开　10. 75 印张　260000 字
2014 年 3 月第 1 版　2014 年 3 月第 1 次印刷
ISBN 978－7－5141－4434－5　定价：28. 00 元
（图书出现印装问题，本社负责调换。电话：010－88191502）

目　　录

CONTENTS

连续调查的差值估计法

张荷观①

摘　要：本文讨论了连续调查时前期与后期总体均值变化的估计，提出了前期与后期总体均值变化的最优组合估计。并给出了估计量的偏差和均方误差，以及均方误差的估计。

关键词：连续调查　前期与后期的变化　组合估计

Difference Estimates in Continuous Surveys

ZHANG Heguan

Abstract. In this paper, the estimate of the change between population mean of earlier stage and population mean of later stage was discussed. The best combined estimation of the change between population mean of earlier stage and population mean of later stage was presented. At the same time, the bias and mean square errors (MSE) of the estimates, and the estimation of the MSE were given.

Key words. continuous surveys, change between earlier stage and later stage, combined estimates

一、引　　言

对同一总体在两个不同时期进行抽样调查时，前期的标志值记为（X_1，X_2，…，X_N），而现期（或后期）的标志值记为（Y_1，Y_2，…，Y_N）。设前期抽取样本量为 n_1 的简单随机样本，而现期则分别从 n_1 和 $N-n_1$、中抽取样本量为 m 和 u_2 的简单随机样本。即 m 是拼配样本量，u_1 和 u_2 则分别是前期和现期的非拼配样本量。于是现期的样本量 $n_2=m+u_2$，而前期的样本量则为 $n_1=m+u_1$。记：

$$\bar{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i,\ \bar{X}=\frac{1}{N}\sum_{i=1}^{N}X_i$$

$$\bar{x}_m=\frac{1}{m}\sum_{i=1}^{m}x_i,\ \bar{x}_u=\frac{1}{u_1}\sum_{i=1}^{u_1}x_i,\ \bar{x}=\frac{m\bar{x}_m+u_1\bar{x}_u}{n_1}$$

$$\bar{y}_m=\frac{1}{m}\sum_{i=1}^{m}y_i,\ \bar{y}_u=\frac{1}{u_2}\sum_{i=1}^{u_2}y_i,\ \bar{y}=\frac{m\bar{y}_m+u_2\bar{y}_u}{n_2}\qquad(1)$$

① **张荷观**：江南大学商学院教授，研究方向：抽样理论，计量经济学。

$$b=\frac{\sum_{i=1}^{m}(x_i-\bar{x}_m)(y_i-\bar{y}_m)}{\sum_{i=1}^{m}(x_i-\bar{x}_m)^2},\ \beta=\frac{\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{N}(X_i-\bar{X})^2}$$

由于

$$\bar{y}_{lr}=\bar{y}_m+b(\bar{x}-\bar{x}_m) \tag{2}$$

和 $\bar{y}_u$ 都可作为现期总体均值 $\bar{Y}$ 的估计，目前常采用组合估计[1, 2]

$$\bar{y}_{tlr}=\lambda\,\bar{y}_{lr}+(1-\lambda)\bar{y}_u \tag{3}$$

作为 $\bar{Y}$ 的估计，其中 λ 应使 $E(\bar{y}_{tlr}-\bar{Y})^2$ 达到最小。

抽样理论目前主要研究现期总体均值 $\bar{Y}$ 的估计，然而对国民经济和社会发展情况的连续调查，不仅需要对现期的总体均值进行估计，同时也需要估计现期与前期总体均值的变化。记

$$D=\bar{Y}-\bar{X} \tag{4}$$

那么现期与前期总体均值之差 D 可以反映现期与前期总体均值的变化。本文把 D 的估计

$$\begin{aligned}\hat{D}&=\bar{y}_{tlr}-\bar{x}\\&=\lambda(\bar{y}_{lr}-\bar{x})+(1-\lambda)(\bar{y}_u-\bar{x})\end{aligned} \tag{5}$$

称为差值估计。由于 $\bar{y}_{lr}$ 和 $\bar{y}_u$ 都可作为 $\bar{Y}$ 的估计，而 $\bar{x}$ 又可作为 $\bar{X}$ 的估计，从而 $\bar{y}_{lr}-\bar{x}$ 和 $\bar{y}_u-\bar{x}$ 都可作为 D 的估计。所以差值估计量 $\hat{D}$ 也是组合估计，但式（5）中的 λ 则应使 $E(\hat{D}-D)^2$ 达到最小。本文给出了 $\hat{D}$ 的偏差和均方误差，以及均方误差的估计。而根据式（5），这又需要先给出 $\bar{y}_{tlr}$ 的偏差和均方误差，以及均方误差的估计。

二、现期总体均值的估计

引理 1　设总体在两个不同时期进行抽样调查时，前期抽取样本量为 n_1 的简单随机样本，而现期则分别从 n_1 和 $N-n_1$ 中抽取样本量为 m 和 u_2 的简单随机样本，则

（i）$$E(\bar{x}_m)=\bar{X},\ Var(\bar{x}_m)=\left(\frac{1}{m}-\frac{1}{N}\right)S_x^2 \tag{6}$$

（ii）$$E(\bar{x}_u)=\bar{X},\ Var(\bar{x}_u)=\left(\frac{1}{u_1}-\frac{1}{N}\right)S_x^2 \tag{7}$$

（iii）$$E(\bar{y}_m)=\bar{Y},\ Var(\bar{y}_m)=\left(\frac{1}{m}-\frac{1}{N}\right)S_y^2 \tag{8}$$

（iv）$$E(\bar{y}_u)=\bar{Y},\ Var(\bar{y}_u)=\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2 \tag{9}$$

其中

$$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2\ ,\ S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$$

证：（ⅰ）由于

$$E(\overline{x}_m | n_1) = \overline{x}$$

则

$$E(\overline{x}_m) = E[E(\overline{x}_m | n_1)] = E(\overline{x}) = \overline{X}$$

记

$$s_x^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(x_i - \overline{x})^2$$

由于

$$\begin{aligned} E[(\overline{x}_m - \overline{X})^2 | n_1] &= E\{[(\overline{x}_m - \overline{x}) + (\overline{x} - \overline{X})]^2 | n_1\} \\ &= \left(\frac{1}{m} - \frac{1}{n_1}\right)s_x^2 + (\overline{x} - \overline{X})^2 \end{aligned}$$

于是

$$\begin{aligned} Var(\overline{x}_m) = E\{E[(\overline{x}_m - \overline{X})^2 | n_1]\} &= E\left[\left(\frac{1}{m} - \frac{1}{n_1}\right)s_x^2 + (\overline{x} - \overline{X})^2\right] \\ &= \left(\frac{1}{m} - \frac{1}{n_1}\right)S_x^2 + \left(\frac{1}{n_1} - \frac{1}{N}\right)S_x^2 \\ &= \left(\frac{1}{m} - \frac{1}{N}\right)S_x^2 \end{aligned}$$

而（ⅱ）、（ⅲ）、（ⅳ）的证明与（ⅰ）类似。

引理2　在引理1的条件下，则

（ⅰ）
$$E(\overline{x}_m - \overline{X})(\overline{x}_u - \overline{Y}) = -\frac{1}{N}S_x^2 \tag{10}$$

（ⅱ）
$$E(\overline{x}_u - \overline{X})(\overline{y}_m - \overline{Y}) = -\frac{1}{N}S_{xy} \tag{11}$$

其中

$$S_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})$$

证：（ⅰ）记

$$\overline{x}_{N-m} = \frac{1}{N-m}\sum_{i=1}^{N-m} x_i$$

对样本量为 $n_1 = m + u_1$ 的简单随机样本，可视为先从 N 中抽取样本量为 m 的简单随机样本，然后再从 $N-m$ 中抽取样本量为 u_1 的简单随机样本。则

$$\begin{aligned} E[(\overline{x}_m - \overline{X})(\overline{x}_u - \overline{X}) | m] &= (\overline{x}_m - \overline{X})E[(\overline{x}_u - \overline{X}) | m] \\ &= (\overline{x}_m - \overline{X})(\overline{x}_{N-m} - \overline{X}) \\ &= -\frac{m}{N-m}(\overline{x}_m - \overline{X})^2 \end{aligned}$$

得

$$E(\bar{x}_m-\bar{X})(\bar{x}_u-\bar{Y})=-\frac{m}{N-m}E(\bar{x}_m-\bar{X})^2=-\frac{1}{N}S_x^2$$

而（ⅱ）的证明（ⅰ）类似。

定理1　设总体在两个不同时期进行抽样调查时，前期抽取样本量为 n_1 的简单随机样本，而现期则分别从 n_1 和 $N-n_1$ 中抽取样本量为 m 和 u_2 的简单随机样本，则

$$E(\bar{y}_{tlr})=\bar{Y}+O\left(\frac{1}{m}\right) \tag{12}$$

证：由式（3）

$$E(\bar{y}_{tlr})=\lambda E(\bar{y}_{lr})+(1-\lambda)E(\bar{y}_u)$$

而由式（2）

$$\bar{y}_{lr}=\bar{y}_m+\frac{u_1}{n_1}(b-\beta)(\bar{x}_u-\bar{X})-\frac{u_1}{n_1}(b-\beta)(\bar{x}_m-\bar{X})$$
$$+\frac{u_1}{n_1}\beta(\bar{x}_u-\bar{X})-\frac{u_1}{n_1}\beta(\bar{x}_m-\bar{X}) \tag{13}$$

根据回归估计理论[3]

$$E(b-\beta)^2=O\left(\frac{1}{m}\right)$$

由引理1得

$$E(\bar{x}_u-\bar{X})^2=O\left(\frac{1}{u_1}\right),\ E(\bar{x}_m-\bar{X})^2=O\left(\frac{1}{m}\right)$$

于是由式（13）得

$$|E(\bar{y}_{lr})-\bar{Y}|\leqslant\frac{u_1}{n_1}E|(b-\beta)(\bar{x}_u-\bar{X})|+\frac{u_1}{n_1}E|(b-\beta)(\bar{x}_m-\bar{X})|$$
$$\leqslant\frac{u_1}{n_1}\sqrt{E(b-\beta)^2}\sqrt{E(\bar{x}_u-\bar{X})^2}+\frac{u_1}{n_1}\sqrt{E(b-\beta)^2}\sqrt{E(\bar{x}_m-\bar{X})^2}$$
$$\leqslant O\left(\frac{1}{m}\right) \tag{14}$$

即 $E(\bar{y}_{lr})=\bar{Y}+O\left(\frac{1}{m}\right)$。而根据引理1，$E(\bar{y}_u)=\bar{Y}$。所以

$$E(\bar{y}_{tlr})=\lambda\left[\bar{Y}+O\left(\frac{1}{m}\right)\right]+(1-\lambda)\bar{Y}$$
$$=\bar{Y}+O(\frac{1}{m})$$

定理2　在定理1条件下，$\bar{y}_{tlr}$的均方误差为

$$MSE(\bar{y}_{tlr})=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2\right]$$
$$+(1-\lambda)^2\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+O\left(\frac{1}{m^2}\right) \tag{15}$$

其中$\rho=\dfrac{S_{xy}}{S_xS_Y}$。

证：根据样本的抽取方法，$\bar{y}_u$ 不仅与 $\bar{y}_m$ 独立，同时与 $\bar{x}_m$ 和 $\bar{x}_u$ 也独立，从而 $\bar{y}_u$ 与 $\bar{y}_{lr}$独立。则根据式（3）

$$E(\bar{y}_{tlr}-\bar{Y})^2=\lambda^2E(\bar{y}_{lr}-\bar{Y})^2+(1-\lambda)^2E(\bar{y}_u-\bar{Y})^2 \tag{16}$$

记 $\bar{y}'=\bar{y}_m-\beta(\bar{x}_m-\bar{X})$，由（2）式得

$$\begin{aligned}E(\bar{y}_{lr}-\bar{Y})^2 &=E[(\bar{y}'-\bar{Y})+\beta(\bar{x}-\bar{X})-(b-\beta)(\bar{x}_m-\bar{x})]^2\\&=E(\bar{y}'-\bar{Y})^2+\beta^2E(\bar{x}-\bar{X})^2+E(b-\beta)^2(\bar{x}_m-\bar{x})^2\\&\quad+2\beta E(\bar{x}-\bar{X})(\bar{y}'-\bar{Y})-2E(b-\beta)(\bar{x}_m-\bar{x})(\bar{y}'-\bar{Y})\\&\quad-2\beta E(b-\beta)(\bar{x}-\bar{X})(\bar{x}_m-\bar{x})\end{aligned} \tag{17}$$

根据引理 1 和引理 2 可得

$$E(\bar{y}'-\bar{Y})^2=\left(\frac{1}{m}-\frac{1}{N}\right)S_y^2(1-\rho^2)$$

$$\beta^2E(\bar{x}-\bar{X})^2=\beta^2\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2=\left(\frac{1}{n_1}-\frac{1}{N}\right)\rho^2S_y^2$$

$$E(\bar{x}-\bar{X})(\bar{y}'-\bar{Y})=0$$

$$E(b-\beta)^2(\bar{x}_m-\bar{x})^2=O\left(\frac{1}{m^2}\right) \tag{18}$$

$$E(b-\beta)(\bar{x}_m-\bar{x})(\bar{y}'-\bar{Y})=O\left(\frac{1}{m^2}\right)$$

$$E(b-\beta)(\bar{x}-\bar{X})(\bar{x}_m-\bar{x})=O\left(\frac{1}{m^2}\right)$$

由式（17）和式（18）得

$$E(\bar{y}_{lr}-\bar{Y})^2=\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2+O\left(\frac{1}{m^2}\right) \tag{19}$$

再根据引理 1

$$E(\bar{y}_u-\bar{Y})^2=\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2$$

于是根据式（16）即得

$$E(\bar{y}_{tlr}-\bar{Y})^2=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2\right]$$
$$+(1-\lambda)^2\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+O\left(\frac{1}{m^2}\right)$$

并且，根据式（16）取

$$\lambda=\frac{E(\bar{y}_u-\bar{Y})^2}{E(\bar{y}_{lr}-\bar{Y})^2+E(\bar{y}_u-\bar{Y})^2} \tag{20}$$

则使 $MSE(\bar{y}_{tlr})$ 达到最小。

定理 3　在定理 1 条件下，则

$$E[mse(\bar{y}_{tlr})]=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2\right]$$

$$+(1-\lambda)^2\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+O\left(\frac{1}{m^2}\right) \tag{21}$$

其中

$$mse(\bar{y}_{tlr})=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)s_{my}^2(1-r^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)s_y^2\right]$$

$$+(1-\lambda)^2\left(\frac{1}{u_2}-\frac{1}{N}\right)s_y^2$$

$$s_{my}^2=\frac{1}{m-1}\sum_{i=1}^{m}(y_i-\bar{y}_m)^2,\ s_{uy}^2=\frac{1}{u_2-1}\sum_{i=1}^{u_2}(y_i-\bar{y}_u)^2,\ s_y^2=\frac{ms_{my}^2+u_2s_{uy}^2}{n_2} \tag{22}$$

$$r=\frac{\sum_{i=1}^{m}(x_i-\bar{x}_m)(y_i-\bar{y}_m)}{\sqrt{\sum_{i=1}^{m}(x_i-\bar{x}_m)^2}\sqrt{\sum_{i=1}^{m}(y_i-\bar{y}_m)^2}}$$

证：根据回归估计理论[4]

$$E(r^2s_{my}^2)=\rho^2S_y^2+O\left(\frac{1}{m}\right)$$

记

$$s_{n_1}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(y_i-\bar{y})^2$$

$$\bar{y}_{N-n_1}=\frac{1}{N-n_1}\sum_{i=1}^{N-n_1}y_i,\ s_{N-n_1}^2=\frac{1}{N-n_1-1}\sum_{i=1}^{N-n_1}(y_i-\bar{y}_{N-n_1})^2$$

由于

$$E[E(s_{my}^2|n_1)]=E(s_{n_1}^2)=S_y^2,\ E[E(s_{uy}^2|n_1)]=E(s_{N-n_1}^2)=S_y^2$$

即

$$E(s_{my}^2)=S_y^2,\ E(s_{uy}^2)=S_y^2$$

于是

$$E[mse(\bar{y}_{tlr})]=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2\right]$$

$$+(1-\lambda)^2\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+O\left(\frac{1}{m^2}\right)$$

在实际调查时，由定理3按式（20）可得λ的近似值为

$$\lambda=\frac{\mathrm{var}(\bar{y}_u)}{mse(\bar{y}_{lr})+\mathrm{var}(\bar{y}_u)} \tag{23}$$

其中

$$mse(\bar{y}_{lr})=\left(\frac{1}{m}-\frac{1}{n_1}\right)s_{my}^2(1-r^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)s_y^2,\ \mathrm{var}(\bar{y}_u)=\left(\frac{1}{u_2}-\frac{1}{N}\right)s_y^2$$

并且当m较大时，$\bar{Y}$的置信度为$1-\alpha$的区间估计近似值为

$$\left(\bar{y}_{tlr}-u_{\frac{\alpha}{2}}\sqrt{mse(\bar{y}_{tlr})},\ \bar{y}_{tlr}+u_{\frac{\alpha}{2}}\sqrt{mse(\bar{y}_{tlr})}\right)$$

其中 $u_{\frac{\alpha}{2}}$ 为标准正态分布的双侧 α 分位数。

三、现期与前期总体均值之差的估计

定理4　在定理1条件下，则

$$E(\hat{D})=D+O\left(\frac{1}{m}\right) \tag{24}$$

证：由式（5）

$$E(\hat{D})=\lambda E(\bar{y}_{lr}-\bar{x})+(1-\lambda)E(\bar{y}_u-\bar{x})$$

根据引理1 $E(\bar{x})=\bar{X}$，则由（14）式即得

$$E(\hat{D})=\lambda\left[D+O\left(\frac{1}{m}\right)\right]+(1-\lambda)D=D+O\left(\frac{1}{m}\right)$$

定理5　在定理1条件下，$\hat{D}$的均方误差为

$$\begin{aligned}MSE(\hat{D})=&\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]\\&+(1-\lambda)^2\left[\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]\\&+2\lambda\left[\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}\right]+O\left(\frac{1}{m^2}\right)\end{aligned} \tag{25}$$

证：由式（5）

$$\begin{aligned}MSE(\hat{D})=&\lambda^2E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})]^2+(1-\lambda)^2E[(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]^2\\&+2\lambda(1-\lambda)E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})][(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]\end{aligned} \tag{26}$$

根据式（19）和引理1与引理2

$$\begin{aligned}E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})]^2&=E(\bar{y}_{lr}-\bar{Y})^2+E(\bar{x}-\bar{X})^2-2E(\bar{x}-\bar{X})(\bar{y}_{lr}-\bar{Y})\\&=\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2\\&\quad+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2-2\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}+O\left(\frac{1}{m^2}\right)\end{aligned}$$

$$\begin{aligned}E[(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]^2&=E(\bar{y}_u-\bar{Y})^2+E(\bar{x}-\bar{X})^2\\&=\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\end{aligned} \tag{27}$$

$$\begin{aligned}&E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})][(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]^2\\&=-E(\bar{x}-\bar{X})(\bar{y}_{lr}-\bar{Y})+E(\bar{x}-\bar{X})^2\\&=-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2+O\left(\frac{1}{m^2}\right)\end{aligned}$$

把式（27）代入式（26）即得

$$MSE(\hat{D})=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$

$$+(1-\lambda)^2\left[\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$
$$+2\lambda\left[\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}\right]+O\left(\frac{1}{m^2}\right)$$

记

$$A=E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})]^2,\ B=E[(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]^2$$
$$C=E[(\bar{y}_{lr}-\bar{Y})-(\bar{x}-\bar{X})][(\bar{y}_u-\bar{Y})-(\bar{x}-\bar{X})]$$

根据式（26），取

$$\lambda=\frac{B-C}{A+B-2C} \tag{28}$$

可使 $MSE(\hat{D})$ 达到最小。

定理 6　在定理 1 条件下，则

$$E[mse(\hat{D})]=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$
$$+(1-\lambda)^2\left[\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$
$$+2\lambda\left[\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}\right]+O\left(\frac{1}{m^2}\right) \tag{29}$$

其中

$$mse(\hat{D})=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)s_{my}^2(1-r^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)s_y^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)s_x^2\right]$$
$$+(1-\lambda)^2\left[\left(\frac{1}{u_2}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)s_x^2\right]$$
$$+2\lambda\left[\left(\frac{1}{n_1}-\frac{1}{N}\right)s_x^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)s_{xy}\right] \tag{30}$$
$$s_{xy}=\frac{1}{m-1}\sum_{i=1}^{m}(x_i-\bar{x}_m)(y_i-\bar{y}_m)$$

证：由于

$$E(s_x^2)=S_x^2,\ E(s_{xy})=S_{xy}$$

则根据定理 3 即得

$$E[mse(\hat{D})]=\lambda^2\left[\left(\frac{1}{m}-\frac{1}{n_1}\right)S_y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_y^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$
$$+(1-\lambda)^2\left[\left(\frac{1}{u_2}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2\right]$$
$$+2\lambda\left[\left(\frac{1}{n_1}-\frac{1}{N}\right)S_x^2-\left(\frac{1}{n_1}-\frac{1}{N}\right)S_{xy}\right]+O\left(\frac{1}{m^2}\right)$$

在实际调查时，根据定理 6 按式（28）可得 λ 的近似值为

$$\lambda=\frac{\mathrm{var}(\bar{y}_u)+\mathrm{cov}(\bar{x},\bar{y}_{lr})}{mse(\bar{y}_{lr})+\mathrm{var}(\bar{y}_u)} \tag{31}$$

其中

$$\operatorname{cov}(\bar{x},\ \bar{y}_{lr}) = \left(\frac{1}{n_1} - \frac{1}{N}\right)s_{xy}$$

国民经济和社会发展情况的连续调查，主要目的之一就是判断 $H_0: D = 0$ 是否成立。根据假设检验理论，若 $|\hat{D}| > u_{\frac{\alpha}{2}}\sqrt{mse(\hat{D})}$，则当样本量很大时就可在 α 水平上拒绝 H_0，认为 $D \neq 0$。并且，D 的置信度为 $1-\alpha$ 的区间估计近似为

$$\left(\hat{D} - u_{\frac{\alpha}{2}}\sqrt{mse(\hat{D})},\ \hat{D} + u_{\frac{\alpha}{2}}\sqrt{mse(\hat{D})}\right)$$

例：某高校采用简单随机抽样方法在 2006 年抽取 150 名本科生调查学生的月伙食费，2007 年则在这 150 名学生中抽取 50 名学生外，又另抽取 100 名学生调查月伙食费，即 $m = 50$，$u_1 = u_2 = 100$，$n_1 = n_2 = 150$。其中 $m = 50$ 的调查数据列于表 1（单位：元）。

表 1　　$m = 50$ 的月伙食费调查数据

x	y	x	y	x	y	x	y	x	y
320	350	420	420	450	470	380	390	410	420
400	410	410	430	480	480	390	420	450	450
420	450	370	380	390	410	350	390	390	420
360	370	320	330	420	430	370	380	400	410
310	320	330	320	440	470	390	410	410	430
300	300	310	340	450	460	420	440	390	420
310	340	380	390	410	440	420	450	370	420
350	380	360	390	390	430	450	430	350	390
370	390	380	370	470	490	390	420	370	370
400	410	430	450	470	460	380	390	390	400

由样本数据已求得

$$\bar{x}_m = 389.80,\ \bar{x}_u = 388.70,\ \bar{x} = 389.07$$

$$\bar{y}_m = 406.60,\ \bar{y}_u = 407.50,\ \bar{y} = 407.20$$

$$s_{my}^2 = 1884.1224,\ s_{uy}^2 = 1897.3507,\ s_y^2 = 1892.9413$$

$$s_x^2 = 1983.2765,\ s_{xy} = 1807.4694,\ r = 0.9432,\ b = 0.9274$$

已知 $N = 12500$，试采用差值法给出 $\bar{Y}$ 与 D 的估计，以及相应均方误差的估计。

（ⅰ）$\bar{Y}$ 和 $MSE(\bar{y}_{tlr})$ 的估计

根据式（2）

$$\bar{y}_{lr} = 405.92$$

由于

$$\operatorname{var}(\bar{y}_u) = \left(\frac{1}{u_2} - \frac{1}{N}\right)s_y^2 = 18.7780$$

$$mse(\bar{y}_{lr}) = \left(\frac{1}{m} - \frac{1}{n_1}\right)s_{my}^2(1 - r^2) + \left(\frac{1}{n_1} - \frac{1}{N}\right)s_y^2 = 15.2409$$

则按式（23）得

$$\lambda = 0.552$$

于是由式（3）和式（22）即得 $\bar{Y}$ 和 $MSE(\bar{y}_{tlr})$ 的估计

$$\bar{y}_{tlr} = 406.63,\ mse(\bar{y}_{tlr}) = 8.4128$$

从而 $\bar{Y}$ 的置信度为 0.95 的区间估计近似为

$$(400.95,\ 412.31)$$

（ⅱ）D 和 $MSE(\hat{D})$ 的估计

由于

$$\mathrm{var}(\bar{y}_u) + \mathrm{cov}(\bar{x},\ \bar{y}_{lr}) = 30.6832,\ mse(\bar{y}_{lr}) + \mathrm{var}(\bar{y}_u) = 34.0189$$

则按式（31）得

$$\lambda = 0.902$$

根据式（5）和式（30）即得 D 和 $MSE(\hat{D})$ 的估计

$$\hat{D} = 17.00,\ mse(\hat{D}) = 4.1666$$

由于

$$\hat{D} > 1.96 \times \sqrt{4.1666} = 4.00$$

从而 0.05 水平上拒绝 $H_0: D = 0$，实际上可认为 $D > 0$，即现期的总体均值超过了前期的总体均值。而 D 的置信度为 0.95 的区间估计近似为

$$(13.00,\ 21.00)$$

四、结　　语

对国民经济和社会发展情况的连续调查，通常更关心现期与前期总体均值的变化。本文讨论了连续调查时现期与前期总体均值变化的差值估计法，提出了现期与前期总体均值变化的最优组合估计。这种差值法通过假设检验可以判断现期与前期总体均值是否存在显著变化，并通过区间估计可进一步给出现期与前期总体均值的变化大小。

本文提出的现期与前期总体均值变化的最优组合估计，通过调整系数 λ 可以有效地提高估计精度。例如在现期与前期总体均值变化的估计时仍取 $\lambda = 0.552$，则

$$\hat{D} = 17.56,\ mse(\hat{D}) = 8.3326$$

即通过调整系数 λ 后明显减小了 $mse(\hat{D})$ 值。

参考文献

[1] Cochran W G. *Sampling Techniques* (3rd ed) [M]. New York: John Wiley & Sons, 1977.

[2] 胡健颖、孙山泽:《抽样调查的理论、方法和应用》,北京大学出版社2000年版。

[3] 冯士雍、施锡铨:《抽样调查——理论、方法和实践》,上海科学技术出版社1996年版。

[4] 张荷观:《回归估计量均方误差的近似值及其估计量的偏差》,载《无锡轻工大学学报》2001年第2期。

[5] 张荷观:《广义比率估计的研究》,载《统计学评论》,中国财政经济出版社2008年版。

贸易增加值率背后是企业增加值率[①]

——针对中国加工贸易出口的讨论与测算分析

高敏雪　葛金梅[②]

摘　要：伴随国际产业分工和全球生产链形成，贸易增加值概念提出并正在逐渐成为一国对外贸易管理的目标。本文以贸易增加值率作为实现贸易增加值最大化之内涵式路径的标志，论证应将企业增加值率作为决定贸易增加值率水平的基本因素；然后基于中国规模以上全出口型工业企业数据，初步估算中国加工贸易出口增加值率水平，并尝试通过行业、地区以及企业登记注册类型的分组拆解分析，探索提高我国贸易增加值率水平的着力点。

关键词：贸易增加值　贸易增加值率　企业增加值率　全出口型企业

The Ratio of Value Added to Gross Output Backs up the Ratio of Value Added to Gross Exports: Experiences from China's Processing Trade

GAO Minxue　GE Jinmei

Abstract. With the globalization of production, policymakers are increasingly aware of the importance of value-added trade, a new objective of foreign trade management. Considering the ratio of value added to gross exports (VAX ratio) plays a vital role in maximizing the value-added trade, we elaborate the close linkage between VAX ratio and the ratio of value added to gross output of enterprises. Then, we get preliminary estimates of China's VAX ratio of processing trade using data of above-scale complete-export-oriented industrial enterprises. Finally, we discuss feasible ways to increase VAX ratio in China based on the decomposition of VAX ratio with respect to industries, regions and types of enterprise registration.

Key words. value-added trade, ratio of value added to gross exports, ratio of value added to gross output, complete-export-oriented enterprises

一、引　　言

贸易增加值是近几年对应传统国际贸易总值而提出的新概念。其背景在于，伴

---

①［基金项目］教育部人文社会科学研究规划基金项目“国民经济核算理论方法国际新进展及其对中国适用性研究”（项目编号：11YJA910002）。成文过程中，曾受益于与李静萍副教授的讨论，特此致谢。

② **高敏雪**：中国人民大学应用统计科学研究中心研究员，统计学院教授，国民经济核算研究所所长。
葛金梅：中国人民大学统计学院在读博士生。