抽样调查: 第 8 周作业

蒋翌坤 20307100013

Exercise 4.8: 3

c

利用回归估计，可以得到 $\bar{y}_{\text{reg}} = 118.36$, $SE(\bar{y}_{\text{reg}}) = 4.07$

Exercise 4.8: 11

c

利用回归估计，可以得到 $\hat{t}_{y\text{reg}} = 45201.8$，95% 置信区间为 $[40784.1, 49619.6]$

Exercise 4.8: 17

a

利用简单抽样估计，可以得到，区域内平均铅含量的 95% 置信区间为 $[100.99, 153.01]$，平均铜含量的 95% 置信区间为 $[32.15, 37.85]$

b

利用事后分层估计，并使用 n_h/n 来估计 W_h, 可以得到，区域内平均铅含量的 95% 置信区间为 $[105.09, 148.91]$，平均铜含量的 95% 置信区间为 $[32.77, 37.23]$

事后分层估计得到的置信区间范围比简单抽样估计得到的置信区间范围更小，说明使用事后分层估计得到的精度更高。

Exercise 4.8: 26

a

$$\bar{y}_1 - \bar{y}_{\mathcal{U}_1} = \frac{\bar{u}}{\bar{x}} - \frac{t_u}{t_x} = \frac{1}{\bar{x}}\Big(\bar{u} - \frac{t_u}{t_x}\bar{x}\Big) \approx \frac{1}{\bar{x}_{\mathcal{U}}}\Big(\bar{u} - \frac{t_u}{t_x}\bar{x}\Big)$$

同理，

$$\bar{y}_2 - \bar{y}_{\mathcal{U}_2} \approx \frac{1}{1-\bar{x}_{\mathcal{U}}}\Big\{\bar{y} - \bar{u} - \frac{t_y - t_u}{N - t_x}(1-\bar{x})\Big\}$$

所以，

$$\text{Cov}(\bar{y}_1, \bar{y}_2) = \text{Cov}(\bar{y}_1 - \bar{y}_{\mathcal{U}_1}, \bar{y}_2 - \bar{y}_{\mathcal{U}_2}) \approx \frac{1}{\bar{x}_{\mathcal{U}}(1-\bar{x}_{\mathcal{U}})}\text{Cov}\Big[\Big(\bar{u} - \frac{t_u}{t_x}\bar{x}\Big), \Big\{\bar{y} - \bar{u} - \frac{t_y - t_u}{N - t_x}(1-\bar{x})\Big\}\Big]$$

b

令 $N_1 = t_x, N_2 = N - t_x, t_{y_1} = t_u, t_{y_2} = t_y - t_u$，则，

$$\begin{aligned}
&\mathrm{Cov}\left[\left(\bar{u} - \frac{t_u}{t_x}\bar{x}\right), \left\{\bar{y} - \bar{u} - \frac{t_y - t_u}{N - t_x}(1 - \bar{x})\right\}\right] \\
&= \mathrm{Cov}(\bar{u}, \bar{y}) - \mathrm{Cov}(\bar{u}, \bar{u}) + \frac{t_y - t_u}{N - t_x}\mathrm{Cov}(\bar{u}, \bar{x}) - \frac{t_u}{t_x}\left[\mathrm{Cov}(\bar{x}, \bar{y}) - \mathrm{Cov}(\bar{x}, \bar{u}) + \frac{t_y - t_u}{N - t_x}\mathrm{Cov}(\bar{x}, \bar{x})\right] \\
&= \left(1 - \frac{n}{N}\right)\frac{1}{n}\left[R_{yu}S_uS_y - S_u^2 + \left(\frac{t_{y_2}}{N_2} + \frac{t_{y_1}}{N_1}\right)R_{xu}S_xS_u - \frac{t_{y_1}}{N_1}R_{xy}S_xS_y - \frac{t_{y_1}t_{y_2}}{N_1N_2}S_x^2\right]
\end{aligned}$$

要使 $\mathrm{Cov}\left[\left(\bar{u} - \frac{t_u}{t_x}\bar{x}\right), \left\{\bar{y} - \bar{u} - \frac{t_y - t_u}{N - t_x}(1 - \bar{x})\right\}\right] = 0$

也就是要证明 $R_{yu}S_uS_y - S_u^2 + \left(\frac{t_{y_2}}{N_2} + \frac{t_{y_1}}{N_1}\right)R_{xu}S_xS_u - \frac{t_{y_1}}{N_1}R_{xy}S_xS_y - \frac{t_{y_1}t_{y_2}}{N_1N_2}S_x^2 = 0$

即 $R_{yu}S_uS_y + \left(\frac{t_{y_2}}{N_2} + \frac{t_{y_1}}{N_1}\right)R_{xu}S_xS_u = S_u^2 + \frac{t_{y_1}t_{y_2}}{N_1N_2}S_x^2 + \frac{t_{y_1}}{N_1}R_{xy}S_xS_y$

接下来没有思路了…，应该是将 R_{yu}、R_{xu} 和 S_U^2 进行变换，使得左式等于右式，但是这变换无从下手。

Exercise 4.8: 32

a

$$\mathrm{E}[\hat{\bar{y}}_{\mathrm{diff}}] = \mathrm{E}(\bar{y}) + \bar{x}_{\mathcal{U}} - \mathrm{E}(\bar{x}) = \bar{y}_{\mathcal{U}} + \bar{x}_{\mathcal{U}} - \bar{x}_{\mathcal{U}} = \bar{y}_{\mathcal{U}}$$

b

$$\mathrm{V}[\hat{\bar{y}}_{\mathrm{diff}}] = \frac{1 - f}{n}(S_y^2 - 2B_1S_{yx} + B_1^2S_x^2) = \frac{1 - f}{n}(S_y^2 - 2S_{yx} + S_x^2)$$

c

$$\mathrm{V}[\hat{\bar{y}}_{\mathrm{diff}}] - \mathrm{V}[\bar{y}_{\mathrm{r}}] = \frac{1 - f}{n}(S_y^2 - 2S_{yx} + S_x^2) - \frac{1 - f}{n}(S_y^2 - 2BS_{yx} + B^2S_x^2) = \frac{1 - f}{n}((2B - 2)S_{yx} + (1 - B^2)S_x^2)$$

所以当 $S_{yx}/S_x^2 = R \cdot S_y/S_x > (B^2 - 1)/(2B - 2)$ 时，$\mathrm{V}[\hat{\bar{y}}_{\mathrm{diff}}] > \mathrm{V}[\bar{y}_{\mathrm{r}}]$

Exercise 4.8: 44

a

对于每个商业类型，总卡车行驶里程及 95% 置信区间如下表所示：

Business	$\hat{t}_u$	CI lower	CI higher
−99	$3 \cdot 10^{11}$	$2.95 \cdot 10^{11}$	$3.05 \cdot 10^{11}$
1	$7.99 \cdot 10^{11}$	$7.84 \cdot 10^{11}$	$8.15 \cdot 10^{11}$
2	$1.25 \cdot 10^{11}$	$1.19 \cdot 10^{11}$	$1.3 \cdot 10^{11}$
3	$1.28 \cdot 10^{11}$	$1.24 \cdot 10^{11}$	$1.33 \cdot 10^{11}$
4	$3.41 \cdot 10^{10}$	$3.15 \cdot 10^{10}$	$3.67 \cdot 10^{10}$
5	$2.5 \cdot 10^{10}$	$2.33 \cdot 10^{10}$	$2.67 \cdot 10^{10}$
6	$2.09 \cdot 10^{11}$	$2.04 \cdot 10^{11}$	$2.13 \cdot 10^{11}$
7	$7.56 \cdot 10^{10}$	$7.15 \cdot 10^{10}$	$7.98 \cdot 10^{10}$
8	$1.01 \cdot 10^{11}$	$9.64 \cdot 10^{10}$	$1.05 \cdot 10^{11}$
9	$1.01 \cdot 10^{11}$	$9.68 \cdot 10^{10}$	$1.06 \cdot 10^{11}$
10	$6.15 \cdot 10^{9}$	$5.5 \cdot 10^{9}$	$6.79 \cdot 10^{9}$
11	$6.96 \cdot 10^{10}$	$6.64 \cdot 10^{10}$	$7.27 \cdot 10^{10}$
12	$2.82 \cdot 10^{9}$	$2.35 \cdot 10^{9}$	$3.3 \cdot 10^{9}$
13	$2.76 \cdot 10^{10}$	$2.51 \cdot 10^{10}$	$3.01 \cdot 10^{10}$
14	$3.01 \cdot 10^{10}$	$2.83 \cdot 10^{10}$	$3.18 \cdot 10^{10}$

附录

解答题目所使用的代码及输出请见：`https://thisiskunmeng.github.io/sampling/hw8.html`