STK4600

Solutions to Ex #13

1. HH = Household

a) $\pi_1 = P(\text{HH1 at 1. draw}) + P(\text{HH1 at 2. draw})$

$= \frac{3}{5} + P(\text{HH2 or 3 at 1. draw}) \cdot P(\text{HH1 at 2. draw} \mid \text{HH2 or 3 at 1.})$

$= \frac{3}{5} + \frac{2}{5} \cdot \frac{3}{4} = \frac{12+6}{20} = \frac{18}{20} = \frac{9}{10}.$

$\pi_2 = \pi_3$ and $\pi_1 + \pi_2 + \pi_3 = 2 \Rightarrow \pi_2 = \pi_3 = \frac{11}{20}$

b) $s = (1,2)$: $\hat{t}_{HT} = \frac{y_1}{\pi_1} + \frac{y_2}{\pi_2} = 1.220.200$

$\hat{t}_R = X_0 \cdot \sum_s y_i / \sum_s x_i = 1.150.000$

$s = (1,3)$: $\hat{t}_{HT} = 1.365.656$

$\hat{t}_R = 1.250.000$

$s = (2,3)$: $\hat{t}_{HT} = 1.163.640$

$\hat{t}_R = 1.600.000$

c) $p(\{1,2\}) = P(\text{not selecting HH3}) = 1 - \pi_3 = 9/20$

$p(\{1,3\}) = P(\text{not selecting HH2}) = 1 - \pi_2 = 9/20$

Since $\sum_s p(s) = 1$: $p(\{2,3\}) = 1 - \frac{18}{20} = 2/20$

or directly:

$p(\{1,2\}) = P(\text{HH1 at 1. draw, HH2 at 2. draw})$

$+ P(\text{HH2 at 1. draw, HH1 at 2. draw})$

$= \frac{3}{5} \cdot \frac{1}{2} + \frac{1}{5} \cdot \frac{3}{4} = 9/20$

Exactly the same:

$p(\{1,3\}) = 9/20$

d) Possible values and corresponding probabilities

s	$\hat{t}_{HT}$	$\hat{t}_R$	$p(s)$
$\{1,2\}$	1 220 200	1.150.000	9/20
$\{1,3\}$	1365656	1.250.000	9/20
$\{2,3\}$	1163640	1.600.000	2/20.

$\Rightarrow E(\hat{t}_{HT}) = \cancel{1.220.200 \times \frac{9}{20} + 1.365.656 \times \frac{9}{20} + 1.163.640 \cdot \frac{2}{20}}$

$= 1.280.000 = t$

Similar: $E(\hat{t}_R) = 1.240.000$

underestimates t on the average

$Var(\hat{t}_{HT}) = E(\hat{t}_{HT} - t)^2 = \sum_s (\hat{t}_{HT}(s) - t)^2 p(s) = \cancel{783.101.326}$ 6.264.810.6

$SE(\hat{t}_{HT}) = \sqrt{Var(\hat{t}_{HT})} = \cancel{27984}$ $\underline{79151}$

$Var(\hat{t}_R) = E(\hat{t}_R - E\hat{t}_R)^2 = \sum_s (\hat{t}_R(s) - 1240.000)^2 p(s) = 1,665 \times 10^{10}$

$\Rightarrow SE(\hat{t}_R) = \sqrt{Var(\hat{t}_R)} = \underline{129035}$

Prefer $\hat{t}_{HT}$

2 a) $\hat{t}_w = N \cdot \frac{y_i/\pi_i}{1/\pi_i} = N y_i = \begin{cases} 10c & \text{if } i = 1, \dots 9 \\ 20c & \text{if } i = 10 \end{cases}$

$\hat{t}_{HT} = y_i/\pi_i = \begin{cases} 9,1c & i = 1, \dots, 9 \\ 200c & i = 10 \end{cases}$

$\Rightarrow$ Obviously choose $\hat{t}_w$!

b) $E(\hat{t}_{HT}) = t$, $E(\hat{t}_w) = 10c \times .11 \times 9 + 20c \times .01 = 10,1c$

slightly biased

$Var(\hat{t}_{HT}) = \sum_{i=1}^{10} (\hat{t}_{HT}(i) - t)^2 \pi_i = 9 \cdot (1.9c)^2 \times .11 + (189c)^2 \times .01$

$= 360,7839 \Rightarrow SE(\hat{t}_{HT}) = 18,99c$

$MSE(\hat{t}_w)$ ~~$Var(\hat{t}_w)$~~ $= 9 \times (1c)^2 \times .11 + (9c)^2 \times .01 = 1.8c^2$

$Var(\hat{t}_w) = 9 \times (0.1c)^2 \times .11 + (9,9c)^2 \times .01 = ,98c^2$, $\underline{SE(\hat{t}_w) = ,99c}$

3. $$E(\hat{t}_{eleph}) = 3\times.05 + 6\times.9 + 12\times.05 = 6.15$$

$$Var(\hat{t}_{eleph}) = (3-6.15)^2\times.05 + (6-6.15)^2\times.9 + (12-6.15)^2\times.05$$
$$= 2.2275$$

4.

$$Var(Z_i) = E(Z_i^2) - (EZ_i)^2$$
$$= EZ_i - (EZ_i)^2 = \pi_i - \pi_i^2 = \pi_i(1-\pi_i)$$

$$Cov(Z_i, Z_j) = E(Z_iZ_j) - \pi_i\pi_j$$
$$= P(Z_iZ_j=1) - \pi_i\pi_j = P(Z_i=1, Z_j=1) - \pi_i\pi_j$$
$$= \pi_{ij} - \pi_i\pi_j$$

b)

$$Var(\hat{t}_{HT}) = \sum_{i=1}^{N}\left(\frac{y_i}{\pi_i}\right)^2 \underbrace{Var(Z_i)}_{\pi_i(1-\pi_i)} + \sum_{i\neq j}\frac{y_i}{\pi_i}\cdot\frac{y_j}{\pi_j}\cdot\underbrace{Cov(Z_i,Z_j)}_{\pi_{ij}-\pi_i\pi_j}$$

$$= \sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i}y_i^2 + \underbrace{\sum_{i=1}^{N}\sum_{j\neq i}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}y_iy_j}_{2\sum_{i=1}^{N}\sum_{j>i} \text{ — '' — }}$$

c)

$$\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(\pi_i\pi_j - \pi_{ij})\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$$

$$= \left.\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\pi_i\pi_j-\pi_{ij}}{\pi_i^2}y_i^2 + \sum_{i=1}^{N}\sum_{j=i+1}\frac{\pi_i\pi_j-\pi_{ij}}{\pi_j^2}y_j^2\right] A$$

$$- 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\pi_i\pi_j-\pi_{ij}}{\pi_i\pi_j}y_iy_j$$

It remains to show: $\underbrace{\text{sum of first 2 terms}}_{A} = \sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i}y_i^2$:

$$A = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(\pi_i\pi_j - \pi_{ij})\frac{y_i^2}{\pi_i^2} + \sum_{j=2}^{N}\sum_{i=1}^{j-1}(\pi_i\pi_j - \pi_{ij})\frac{y_j^2}{\pi_j^2}$$

$$= \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} (\pi_i\pi_j - \pi_{ij}) \frac{y_i^2}{\pi_i^2} = \sum_{i=1}^{N} \frac{y_i^2}{\pi_i^2} \underbrace{\sum_{\substack{j=1 \\ j\neq i}}^{N} (\pi_i\pi_j - \pi_{ij})}_{\pi_i(1-\pi_i)}$$

$$= \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i} y_i^2$$

5\. a) $\pi_1 = p(\{1,4\}) = 0.2$

$\pi_2 = p(\{2,4\}) = 0.3$

$\pi_3 = p(\{3,4\}) = 0.5$

$\pi_4 = 1$

$\pi_{14} = 0.2$, $\pi_{24} = .3$, $\pi_{34} = .5$

all other $\pi_{ij} = 0$: $\pi_{12} = \pi_{13} = \pi_{23} = 0$

b) $s = \{i, 4\}$

$$\hat{V}(\hat{t}_{\pi r}) = \frac{\pi_i\pi_4 - \pi_{i4}}{\pi_{i4}} \left(\frac{y_i}{\pi_i} - \frac{y_4}{\pi_4}\right)^2 = 0\ !$$

since $\pi_i\pi_4 = \pi_i = \pi_{i4}$ for $i = 1, 2, 3$

Variance estimate is meaningless