0070-WDZ

1) $N=20$ $n=4$ OSUuå

$y=$ ekonomiska tillgångar i tkr
$m=$ medlemsantal

a) $\hat{t}=N\cdot\bar{y}$

$\bar{y}=\frac{1}{n}\sum_{k\in s}y_k=\frac{16+96+3,6+0,8}{4}=29,1$

$\hat{t}_y=N\cdot\bar{y}=20\cdot 29,1=582$

$s_y^2=\frac{\sum_{k\in s}y_k^2-n\bar{y}^2}{n-1}=\frac{9485,6-4\cdot 29,1^2}{4-1}=$

$=\frac{9485,6-3387,24}{3}=2032,786667$

$\hat{V}(\hat{t})=N^2\cdot\hat{V}(\bar{y})=N^2\cdot\left(1-\frac{n}{N}\right)\frac{s^2}{n}=$

$=20^2\cdot\left(1-\frac{4}{20}\right)\frac{2032,786667}{4}=$

$400\cdot 0,8\cdot 508,1966668=162622,9334$

95% KI $\alpha=0,05$ $z_{0,05/2}=1,96$

$\hat{t}_y\pm z_{\alpha/2}\cdot\sqrt{\hat{V}(\hat{t})}=582\pm 1,96\cdot\sqrt{162622,9334}$

$=582\pm 790,4000639$

$\{-208,4;1372,4\}$

1 b) -208,4 & 1372,4 väldigt långt intervall, -208,4 då är föreningarna i skuld. Ingen av föreningarna i stickprovet är i skuld. Blir så pga ett extremvärde åt andra hållet (96).

0070-WDZ

1c) n? $\frac{790,4}{2} = 395,2$ $Z_{0,05/2} = 1,96$

$$D^2 = \frac{B^2}{Z^2_{\alpha/2}} = \frac{395,2^2}{1,96^2} = 40655,72678$$

$$n \geq N^2 \frac{N\sigma_y^2}{D^2(N-1) + N^2\sigma_y^2} =$$

$$= 20^2 \frac{20 \cdot 2032,786667}{40655,73^2(20-1) + 20^2 \cdot 2032,786667}$$

$$= 400 \cdot \frac{40655,73334}{1585573,476} = 10,25641106$$

$$n \geq 10,26$$

Minst 11 foreningar för att ha en felmarginal på 395,2 tkr vid 95% KI.

0070-WDZ

1 d) Skatta M, formlerna för vanligt OSU när man skattar totalen är samma som för klusterurval när man skattar M men använder beteckningar från vanlig OSU.

$\hat{t}_m = N \cdot \bar{m}$

$\bar{m} = \frac{1}{n} \cdot \sum_{k \in s} m_k = \frac{116+72+16+24}{4} = 57$

$\hat{t}_m = 20 \cdot 57 = 1140$

$S_m^2 = \frac{\sum_{k \in s} m_k^2 - n \cdot \bar{m}^2}{n-1} = \frac{19472 - 4 \cdot 57^2}{4-1} = 2158,6667$

$\hat{V}(\hat{t}_m) = N^2 \cdot \left(1 - \frac{n}{N}\right) \frac{S^2}{n} = 20^2 \left(1 - \frac{4}{20}\right) \frac{2158,6667}{4}$

$= 400 \cdot 0,8 \cdot 539,6667 = 172693,336$

95% KI för $\hat{t}_m$ $\quad Z_{0,05/2} = 1,96$

$\hat{t}_m \pm Z_{\alpha/2} \cdot \sqrt{\hat{V}(\hat{t}_m)} = 1140 \pm 1,96 \cdot \sqrt{172693,336}$

$= 1140 \pm 814,5051992$

0070-WDZ

1e) Nej. I undersökningen är vi intresserade av föreningarna inte av der enskilda medlemmaarna.
Våra objekt är alltså föreningarna och inte medlemmarna.
Men man kanske skulle kunna använda antalet medlemmar som en hjälpvariabel för att få bättre precision i punktskattningarna för föreningarna.

0070-WDZ

2/ $N_1=15000 \quad N_2=15000 \quad N_3=20000$

$n=3000 \quad N_1+N_2+N_3=50000$

a) OSUuå andel $=P$

$\hat{p}=\frac{1}{n}\sum_{k\in s} y_k=\frac{960+600+120}{3000}=0{,}56$

$\hat{V}(\hat{p})=\left(1-\frac{n}{N}\right)\frac{\hat{p}(1-\hat{p})}{n-1}=\left(1-\frac{3000}{50000}\right)\frac{0{,}56\cdot 0{,}44}{2999}=$

$=0{,}94\cdot 0{,}0000821607=0{,}0000772310770$

90% KI för $\hat{p}$ $\quad z_{0{,}1/2}=1{,}6449$

$\hat{p}\pm z_{\alpha/2}\sqrt{\hat{V}(\hat{p})}=0{,}56\pm 1{,}6449\cdot\sqrt{0{,}0000772310 7}$

$=0{,}56\pm 0{,}01445558082$

$\{0{,}54554; 0{,}57446$

b) Skatta P $\quad$ OSU$_{str}$ uå

Stratum	N_k	n_k	W_k N_k/N	$\hat{p}_k$	$W_k\cdot\hat{p}_k$
Öst	15000	1200	0,3	960/1200=0,8	0,24
Mitten	15000	1200	0,3	600/1200=0,5	0,15
Väst	20000	600	0,4	120/600=0,2	0,08
Tot	50000	3000			0,47

$\hat{p}_{str}=\sum_{k=1}^{L} W_k\hat{p}_k=0{,}47$

0070-WDZ

2b forts

Stratum	W_k^2	$1-\frac{n_k}{N_k}$	$\frac{\hat{p}_k(1-\hat{p}_k)}{n_k-1}$	$W_k^2\left(1-\frac{n_k}{N_k}\right)\frac{\hat{p}_k(1-\hat{p}_k)}{n_k-1}$
Öst	0,09	$1-\frac{1200}{15000}=0,92$	0,0001334445	0,00001104920 8
Mitten	0,09	— " — $=0,92$	0,000208507 09	0,0000172643 8691
Väst	0,16	$1-\frac{600}{20000}=0,97$	0,000267111 8531	0,00004144557 96
Tot				0,0000697693 54 26

$$\hat{V}(\hat{p})=\sum_{k=1}^{L} W_k^2\left(1-\frac{n_k}{N_k}\right)\frac{\hat{p}_k(1-\hat{p}_k)}{n_k-1}=0,0000697693 5426$$

90% KI för $\hat{p}_{str}$ $\quad z_{\alpha/2}=1,6449$

$$\hat{p}_{str}\pm z_{\alpha/2}\sqrt{\hat{V}(\hat{p})_{str}}=0,47\pm 1,6449\sqrt{0,0000697693 5426}$$

$$=0,47\pm 0,0137395292 3$$

$$\{0,4562605\,;\,0,4837395\}$$

2c) (Post) Efterstratifiering. Eftersom N_k är kända här så kan jag skatta W_k och de vikterna behöver man för formlerna vid poststratifiering. Alltså man behöver vikta hur stor del av ett stratums ex $\bar{y}$ som ska tas med när man skattar populationens $\bar{y}$. Om N_k inte varit känt hade vi kunnat göra domänskattningar. Eftersom det historiskt →

0070-WD2

2c) forts. sett har skilt sig politiskt mellan områdena, alltså homogent inom området & heterogent mellan områdena i avseende på andelen som röstar på detta parti.
Så är det lämpligt att använda oss av områdena som stratifieringsvariabel
Nu har vi fått ~~ett~~ litet snedfördelning urvalet om man skulle göra proportionell allokering. Bara 600 objekt från det största stratumet. Men genom efterstratifiering får deras resultat väga tyngre i populationsskattningen.
Punktskattningen hade blivit densamma som i b) men felmarginalen hade blivit längre än vid stratifierat och då även standardfelet högre. Pga att n_k blir slumpmässig så blir $\hat{p}$ en kvot mellan 2 slumpvariabler och då blir också variansformlerna annorlunda vilket ger större varians & större standardfel & större felmarginal än vid stratifierat

0070-WD2

3 $N=60$ OSUuå $n=8$

x_i = försäljning i 3mån innan tkr
y_i = försäljning i 3mån efter tkr

$\tau_x = 18000$

a) skatta τ_y med kvotestimator. Alltså inte reg? Borde vara det men det svarar jag väl på i c.

$$\tau_{kvot} = \hat{R}\,\tau_x$$

OSU → $\bar{y} = \frac{\sum_{k \in s} y_k}{n} = \frac{2780}{8} = 347{,}5$

$$\hat{\tau}_y = \overset{N \cdot \bar{y}_s}{60 \cdot 347{,}5} = 20850$$

Från stickprov → $\bar{x} = \frac{\sum_{k \in s} x_k}{n} = \frac{2600}{8} = 325$

$$\hat{\tau}_x = N \cdot \bar{x} = 60 \cdot 325 = 19500$$

$$\hat{R} = \frac{20850}{19500} = 1{,}069230769 \quad \left(\hat{R} = \frac{\hat{\tau}_y}{\hat{\tau}_x}\right)$$

$$\tau_{kvot} = \hat{R}\,\tau_x = 1{,}069230769 \cdot 18000 = 19246{,}15384$$

→

0070-WDZ

3a forts

$$\hat{V}(\hat{\tau}_{kvot}) = N^2 \cdot \left(1-\frac{n}{N}\right) \frac{\sum_{k \in s} y_k^2 - 2\hat{R}\sum_{k \in s} x_k y_k + \hat{R}^2 \sum_{k \in s} x_k^2}{n-1}$$

$$= 60^2 \cdot \left(1-\frac{8}{60}\right) \frac{1059302 - 2 \cdot 1{,}06923 \cdot 815227 + 1{,}06923^2 \cdot \underbrace{931360}}{8-1}$$

$$= 3600 \cdot 0{,}866667 \cdot \frac{1059302 - 1743330{,}33 + 1064779{,}921}{7}$$

$$= -\text{''}- \cdot 54393{,}08443$$

$$= 169706488{,}7$$

Standarfelet för $(\hat{\tau}_{kvot_{ty}}) = \sqrt{169706488{,}7}$

$= 13027{,}1443$

3b) OSUnå

$\hat{\tau}_y$ = från tidigare $= 20850$

$$s_y^2 = \frac{1059302 - 8 \cdot 347{,}5^2}{8-1} = 13321{,}71429$$

$$\hat{V}(\hat{\tau}_y) = N^2 \cdot \left(1-\frac{n}{N}\right)\frac{s_y^2}{n} = 60^2 \cdot \left(1-\frac{8}{60}\right)\frac{13321{,}71429}{8}$$

$$= 5195470{,}571$$

$$s(\hat{\tau}_y) = \sqrt{5195470{,}571} = 2279{,}357491$$

svar

0070-WDZ

3c Sambandet mellan x och y går inte genom origo. Och då ska man inte använda kvotestimatorn utan regressionsskattningen som har ett intercept. Ej rimligt att 0 kampanj ger 0 försäljning.

4 a) Aktualitet & Noggrannhet

Aktualitet: Statistiken ska vara aktuell så inferens "gäller" och inte två år tillbaka.

Noggrannhet: Utfört en välplanerad undersökning, och redovisat för alla möjliga fel som kan uppstå

Målkonflikt: Noggrannhet tar tid, och då kan inferensen inte vara aktuell. Kanske så att man får tumma lite på noggrannheten för att få ut aktuell statistik.

0070-WDZ

4b) Domänskattningar kan man göra när man inte har någon ram för vilka som ingår i den domänen. Man drar ett OSU så kollar man hur många som tillhör den sökta domänen. Och därefter gör punktskattningar på domänen. Hur många n för domänen blir då slump-mässigt, allts n_k slumpmässig. Och detta påverkar variansen och därmed osäkerheten i skattningarna. N_k är också oftast okänt i domänskattningarna vilket påverkar varians & osäkerhet. Jämfört med stratifierat urval så känner vi till N_k (stratumpopulationen). Och bestämmer n_k i förväg vilket ger bättre precision i skattningar och mindre osäkerhet. ~~Här~~ kan vi också gardera oss för sneda urval (alltså att urvalet blir representativ). Det kan vi inte göra i domänskattningar och vi har oftast ingen aning om att det är snett urval.

0070-WD2

4,c Ingen väg ut, kanske borde ha alltid, oftast, ..., aldrig och en vet inte/vill ej uppge

$\hat{p}_B = 0$ $\quad$ $\hat{p}_{min} = W_s \cdot \hat{p}_s = 0{,}2 \cdot 0{,}8 = 0{,}16$

$\hat{p}_B = 1$ $\quad$ $\hat{p}_{max} = W_s \hat{p}_s + W_B = 0{,}2 \, 0{,}8 + 0{,}7 = 0{,}86$

0070-WDZ

$\hat{V}(\bar{y}) = \left(1-\frac{n}{N}\right)\frac{s^2}{n}$

5 $N=5$ $n=2$

$U=\{0,1,2,3,4\}$

a) $\binom{5}{2} = \frac{5!}{2!3!} = \frac{4\cdot5}{1\cdot2} = \frac{20}{2} = 10$ möjliga stickprov

		$\bar{y}$	$s_y^2 = \frac{\sum_{k\in s} y_k^2 - n\bar{y}^2}{n-1}$	$\hat{V}(\bar{y})$
1	0,1	0,5	0,5	0,15
2	0,2	1	2	0,6
3	0,3	1,5	4,5	1,35
4	0,4	2	8	2,4
5	1,2	1,5	0,5	0,15
6	1,3	2	2	0,6
7	1,4	2,5	4,5	1,35
8	2,3	2,5	0,5	0,15
9	2,4	3	2	0,6
10	3,4	3,5	0,5	0,15
				7,5

b) $E(s_y^2) = \overline{s_y^2} = \frac{\sum_{i=1}^{M} s_{i,y}^2}{M} = \frac{25}{10} = 2,5$

$\sigma_y^2 = \frac{\sum_{k\in U} y_k^2 - N\cdot\mu^2}{N} = \frac{30-5\cdot2^2}{5} = 2$

$\Rightarrow \mu = \frac{0+1+2+3+4}{5} = 2$

Nei den är inte vvr. Men om man tar $\frac{N}{N-1}\cdot\sigma_y^2$ så blir det samma resultat.

0070-WDZ

5c) $E(\hat{V}(\bar{y})) = \overline{\hat{V}(\bar{y})} = V(\bar{Y})$

Vet inte riktig vilken du vill att jag ska beräkna

$$\overline{\hat{V}(\bar{y})} = \frac{\sum_{i=1}^{M} \hat{V}(\bar{y})_i}{M} = \frac{\text{från tabell } 7{,}5}{10} = 0{,}75$$

$$V(\bar{Y}) = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n} = \frac{5-2}{5-1} \cdot \frac{2}{2} = 0{,}75$$