

Виконавець: Симонова Вероніка

Умова задачі: Аналіз сімей, які отримують субсидій. В залежності від результатів розглядається питання про підвищення субсидій.

Змінні: 1) Величина субсидій (Т - сумарне значення)

2) Загальний прибуток на 1 сім'ю (У - середнє значення)

3) Оренда квартири (Р - частка сімей, що проживають в орендованому житлі)

Будували просту ВВбп 2 методом ( розрахунок в Exsel)

	Номер блока	Номер домогосподарства
1	9	1
2	11	3
3	23	2
4	23	7
5	39	2
6	43	1
7	52	7
8	53	1
9	59	1
10	59	4

$N < -60 * 8$

$n < -10$

$T < -\text{sum}(\text{st\$`9`})$  # сумарне значення субсидій за 10 господарств

$T1 < -N * T / n$  # оцінка сумарного значення за усім селом

$Y < -\text{mean}(\text{st\$`11`})$  # середнє значення прибутку на одну сім'ю

$P < -1 - \text{sum}(\text{as.numeric}(\text{st\$`14`} < 2)) / 10$  # Р - частка сімей, що проживають в орендованому житлі

Результати:

Т	У	Р
3827184	80464.1	0.1

Оцінки для дисперсії оцінок:

$11 < -\text{mean}(\text{st\$`9`})$  # середнє значення субсидії

$T1 < -N^2 * \text{sum}((\text{st\$`9`} - T1)^2) / (n - 1)$  # дисперсія для субсидії

$Y1 < -1 / (n - 1) * \text{sum}((\text{st\$`11`} - Y)^2)$  # дисперсія для прибутку

$P1 < -P * (1 - P) * n / (n - 1)$  #дисперсія для частки

T1	Y1	P1
1.931034e+13	219901105	0.1

2. Для того, щоб оцінити розмір вибірки, необхідний для знаходження 95% довірчих інтервалів з точністю в 10% для усіх параметрів дослідження одночасно, знайдемо відповідні розміри вибірки для кожного параметра. Тобто, такі  $n$ , що  $P(|(\bar{y} - \bar{y}_n) / \bar{y}_n| \leq 0.1) \approx 0.95$ , для відповідних  $y$ .

$\alpha < -0.05$

$e < -0.1$

$z < -qnorm(1 - \alpha / 2)$

$CV < -c(sqrt(T1) / T2, sqrt(Y1) / Y, sqrt(P1) / P)$

$En < -z^2 * CV^2 / (e^2 + CV^2 * z^2 / N)$

Total	Mean	Proportion
247	13	427

Отже, якщо ми використаємо вибірку з 427 елементів, то зможемо розраховувати на задану точність оцінок.

```
n<-427
N<-480
i<-sample(1:N,n)
alpha<-0.05
z<-qnorm(1-alpha/2)
st1<-data.frame(full_stat$`9`[i],full_stat$`11`[i],full_stat$`14`[i])
colnames(st1)<-c("T","Y","P")
T2<-N*sum(st1$T)/n
Y<-mean(st1$`Y`)
T11<-sum(st1$T)/n
P<-1-sum(as.numeric(st1$P< 2))/n
T1<-N^2*sum((st1$T-T11)^2)/(n-1)
Y1<-sum((st1$Y-Y)^2)/(n-1)
P1<-P*(1-P)*n/(n-1)
q<-c(T1,Y1,P1)
Q<-c(T2,Y,P)
val1<-Q-z*sqrt(q*(1/n-1/N))
val2<-Q+z*sqrt(q*(1/n-1/N))
```

	$\hat{\theta}^-$	$\hat{\theta}$	$\hat{\theta}^+$
Y	83428.42	84489.46	85550.51
P	0.1340824	0.1451991	0.1563157
T	1230482	1297786	1365091