

# ЛАБОРАТОРНА РОБОТА №1

Верцімаха Оксана, прикладна статистика

В роботі досліджуються такі дані населення:

- 1. Величина субсидій (Т - сумарне значення)
- 2. Загальний прибуток сім'ї (Y - середнє значення)
- 3. Оренда квартири (Р - частка сімей, що проживають в орендованому житлі)

```
1. N<-60*8
   n1<-10
   X<-sample(1:N, n)
   Yb1<-X%/%8+1
   Yh<-X%%8
   for(i in 1:n1){if(Yh[i]==0){
     Yb1[i]<-Yb1[i]-1
     Yh[i]<-8
   }
 }
 M<-cbind(Yb1, Yh)
 colnames(M)<-c("block","house")
 rownames(M)<-1:n1

> M
```

	block	house
1	22	3
2	36	5
3	55	2
4	14	6
5	26	7
6	9	4
7	49	8
8	16	1
9	13	2
10	22	4

Тестова вибірка з 10 домогосподарств, обрана за допомогою простого випадкового відбору без повернення, ПВВБП

```
Data<-data.frame(test_sample$V3,test_sample$V11, test_sample$V9,
  strtoi(test_sample$V19), stringsAsFactors = F)
colnames(Data)<-c("Residents", "Income", "Subsidies", "Rent")
```

```
#Estimation of total
(total of government transfer payments received by the population)
t<-N*sum(Data$Subsidies)/n1
```

```
#Estimation of mean (mean total income of a household)
m<-mean(Data$Income)
```

```
#Estimation of proportion
(proportion of families with rented household)
p<- 1 - sum(is.na(Data$Rent))/n1
```

Отримані значення оцінок:

Total, $\hat{t}$	Mean, $\hat{m}$	Proportion, $\hat{p}$
1199280	83505.6	0.3

*# Estimation of variance for the total*

```
tm<-sum(Data$Subsidies)/n1
St<-(N^2)*sum((Data$Subsidies-tm)^2)/(n1-1) #середнє для субсидій
```

*# Estimation of variance for the mean*

```
Sm<-sum((Data$Income-m)^2)/(n1-1)
```

*# Estimation of variance for the proportion*

```
Sp<-p*(1-p)*n1/(n1-1)
```

Отримані значення оцінок для дисперсій оцінок:

Total, $\hat{S}^2$	Mean, $\hat{S}^2$	Proportion, $\hat{S}^2$
1.380693e+12	309924528	0.23

- 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  $P\left(\left|\frac{\bar{y}-\bar{y}_U}{\bar{y}_U}\right| \leq 0.1\right) \approx 0.95$ , для відповідних  $y$ .

```
alpha<-0.05
e<-0.1
z<-qnorm(1-alpha/2)
CV<-c(sqrt(St)/t, sqrt(Sm)/m, sqrt(Sp)/p)
En<-z^2*CV^2/(e^2+CV^2*z^2/N)
```

Total, $n$	Mean, $n$	Proportion, $n$
209	17	324

Отже, якщо ми використаємо вибірку з 324 елементів, то зможемо розраховувати на задану точність оцінок.

### 3.

```
#ПВВБП
n<-324
ind<-sample(1:N, n)

# Завантажуюмо дані
Sample<-data.frame(full_stat$V3[ind],full_stat$V11[ind],
  full_stat$V9[ind],
  strtoi(full_stat$V19[ind]), stringsAsFactors = F)
colnames(Sample)<-c("Residents", "Income", "Subsidies", "Rent")

t<-N*sum(Sample$Subsidies)/n
m<-mean(Sample$Income)
tm<-sum(Sample$Subsidies)/n
p<-1 - sum(is.na(Sample$Rent))/n
St<-(N^2)*sum((Sample$Subsidies-tm)^2)/(n-1)
Sm<-sum((Sample$Income-m)^2)/(n-1)
Sp<-p*(1-p)*n/(n-1)
alpha<-0.05
z<-qnorm(1-alpha/2)
S<-c(St,Sm,Sp)
est<-c(t,m,p)
val1<-est-z*sqrt(S*(1/n-1/N))
val2<-est+z*sqrt(S*(1/n-1/N))
```

	$\hat{\theta}_-$	$\hat{\theta}$	$\hat{\theta}_+$
<b>Total, <math>\hat{t}</math></b>	1254399.4	1390840	1527280.6
<b>Mean, <math>\hat{m}</math></b>	82339.286	84379.281	86419.275
<b>Proportion, <math>\hat{p}</math></b>	0.12896007	0.15123457	0.17350906