

Лабораторна робота 1

В роботі досліджуються: величина субсидій (сумарне), загальний прибуток (середнє), оренда житла (частка орендованого).

Отримані результати:

```
data_test_selected = get_data(N_test)
values_test = Value(data_test_selected)
variances_test = Variance(data_test_selected)
```

Тестова вибірка з 10 домогосподарств

Block	36	17	4	47	49	27	6	54	37	28
House	5	6	1	7	5	1	1	5	1	8

Оцінки величин за тестовою вибіркою

Total	Mean	Proportion
689568.0	81786.1	0.1

Оцінки дисперсій за тестовою вибіркою

Variance of total	Variance of mean	Variance of proportion
242236647403	87778569	0.0098

```
cv = sqrt(variances_test) / values_test
N_conf = ceiling(x^2 * cv^2 / (e^2 + cv^2 * x^2 / N_houses))
N_conf = max(N_conf)
```

Оцінки розміру вибірки

N for total	N for mean	N for proportion
140	5	211

$N_{conf} = 211$ — розмір вибірки, необхідний для побудови 95% довірчих інтервалів

Для заданого $n = 211$ побудуємо довірчі інтервали

```
data_conf_selected = get_data(N_conf)
values_conf = Value(data_conf_selected)
variances_conf = Variance(data_conf_selected)
val1 = values_conf - x * sqrt(variances_conf * (1/N_conf - 1/N_houses))
val2 = values_conf + x * sqrt(variances_conf * (1/N_conf - 1/N_houses))
```

	Total	Mean	Proportion
Оцінка знизу	143901.8	86758.56	0.01464
Наближене значення	167392.3	87165.4	0.01612
Оцінка зверху	190882.7	87572.24	0.01761178

Функції для підрахунку та роботи з даними:

- Формує вибірку потрібної довжини ПВВбП

```
get_data = function(amount){  
  index_selected = sample(1:N_houses, amount, replace=F)  
  data_selected = DATA[index_selected, c(1, 2, 3, 9, 11, 19)]  
  names(data_selected) = c("block", "unit", "residents", "income", "subsidies", "rent")  
  data_selected  
}
```

- Підраховує оцінки для сумарного, середнього та частки за даними

```
Value = function(data_sample){  
  N_sample = nrow(data_sample)  
  #Estimation of total  
  t_subsidies = N_houses * sum(data_test_selected$subsidies) / N_sample  
  #Estimation of Mean  
  m_income = mean(data_test_selected$income)  
  #Estimation of Proportion  
  p_rent = sum(data_test_selected$rent != '.') / N_sample  
  c(t_subsidies, m_income, p_rent)  
}
```

- Підраховує оцінки дисперсій за даними

```
Variance = function(data_sample){  
  N_sample = nrow(data_sample)  
  # Estimation of total subsidies variance  
  var_t_subsidies = N_houses^2 * (1 - N_sample/N_houses) * sd(data_test_selected$subsidies)^2 /  
  N_sample  
  # Estimation of mean variance  
  var_m_income = (1 - N_sample/N_houses) * sd(data_test_selected$income)^2 / N_sample  
  # Estimation of proportion variance  
  p_rent = sum(data_test_selected$rent != '.') / N_sample  
  var_p_rent = p_rent * (1 - p_rent) * (1 - N_sample/N_houses) / (N_sample - 1)  
  c(var_t_subsidies, var_m_income, var_p_rent)  
}
```