

## Лабораторна робота №2

В цій роботі досліджується параметр R як відношення сумарних субсидій до сумарної кількості людей у будинку.

За допомогою ПВВБП створюється вибірка з половини господарств.

```
N_houses = 60 * 8
N_test = N_houses/2
DATA = read.table("/home/asta/Nonparametric-statistics/database.csv", header = FALSE, sep = ",")
get_data = function(amount){
  index_selected = sample(1:N_houses, amount, replace=F)
  data_selected = DATA[index_selected, c(1, 2, col_residents, col_subsidies)]
  data_selected
}
data_test = get_data(N_test)
t_up = sum(data_test[,4])*N_houses/N_test
t_down = sum(data_test[,3])*N_houses/N_test
ratio = t_up/t_down
```

- оцінка відношення сумарних

### Метод Лінеаризації:

```
D_linear = N_houses*(N_houses/N_test-1)*(1/t_down)^2 * (var(data_test[,4])
+ratio^2*var(data_test[,3])-2*ratio*var(data_test[,3], data_test[,4]))
```

### Джек Найф:

```
n_jk = nrow(data_test)
mean_jk = mean(data_test[,4])/mean(data_test[,3])
D_jk = 0
for (i in 1:n_jk) {
  data_jk = data_test[-i,]
  D_jk = D_jk + (sum(data_jk[,4])/sum(data_jk[,3]) - mean_jk)^2
}
D_jk = D_jk * (n_jk - 1)/n_jk
```

### Бутстреп:

```
n_bs = 1000
ratio_bs = rep(0, n_bs)
for (i in 1:n_bs) {
  # data_bs = data_test[sample(1:N_test, runif(1, min=1, max=N_test)), ]
  data_bs = data_test[sample(1:N_test, N_test, replace = TRUE), ]
  ratio_bs[i] = sum(data_bs[,4])/sum(data_bs[,3])
}
#D_bs = var(ratio_bs)
mean_bs = mean(ratio_bs)
D_bs = sum((ratio_bs - mean_bs)^2)/(n_bs - 1)
```

D_linear	D_jk	D_bs
3616.206	7229.705	7341.752

### Довірчі інтервали:

$\alpha = 0.05$

$x = \text{qnorm}(1 - \alpha / 2)$

$\text{variances} = c(D\_linear, D\_jk, D\_bs)$

$\text{val1} = \text{ratio} - x * \text{sqrt}(\text{variances} * (1/N\_test - 1/N\_houses))$

$\text{val2} = \text{ratio} + x * \text{sqrt}(\text{variances} * (1/N\_test - 1/N\_houses))$

### Побудова графіку

`library(ggplot2)`

`d = data.frame(x = c("linearization", "jack_knife", "bootstrap"), y = c(ratio, mean_jk, mean_bs), ylo = val1, yhi = val2)`

`ggplot(d, aes(x = x, y = y)) +`

`geom_point(size = 3) +`

`geom_errorbar(aes(ymin = ylo, ymax = yhi))`

