**Мета**

Дослідити вплив різних характеристик на рівень гемоглобіну людей не молодше 20 років.

**Початкові моделі**

Побудуємо лінійні моделі:

Дані: всі спостереження для яких age Є [20; 60], weight Є [60, 75].

Залежна змінна: hemoglobin

Модель 1:

Незалежна змінна: age

Модель 2:

Незалежна змінна: weight

data\_age\_20\_60\_weight\_60\_75 <- data %>% filter(age >= 20, age <= 60, weight >= 60, weight <= 75)

lm\_in\_age\_out\_hemoglobin <- lm(hemoglobin ~ age, data = data\_age\_20\_60\_weight\_60\_75)

plot\_lm\_in\_age\_out\_hemoglobin <- ggplotRegression(lm\_in\_age\_out\_hemoglobin)

ggsave("./output/Lab3\_Regression/plot\_lm\_in\_age\_out\_hemoglobin.png",

plot = plot\_lm\_in\_age\_out\_hemoglobin)

lm\_in\_weight\_out\_hemoglobin <- lm(hemoglobin ~ weight, data = data\_age\_20\_60\_weight\_60\_75)

plot\_lm\_in\_weight\_out\_hemoglobin <- ggplotRegression(lm\_in\_weight\_out\_hemoglobin)

ggsave("./output/Lab3\_Regression/plot\_lm\_in\_weight\_out\_hemoglobin.png",

plot = plot\_lm\_in\_weight\_out\_hemoglobin)

Обмеження на змінні age та weight накладено з метою полегшення виявлення кореляції.

Візуалізуємо моделі:

A graph with a red line

Description automatically generated

Модель hemoglobin ~ age

A graph of a graph

Description automatically generated

Модель hemoglobin ~ weight

Має місце деяка кореляція.

Побудуємо лінійну модель:

Дані: весь очищений датасет (991284 спостережень)

Залежна змінна: hemoglobin

Незалежні змінні: age, weight

model <- lm(hemoglobin ~ age + weight, data = data)

model\_hc0 <- coeftest(model, vcov. = vcovHC(model, type = "HC0"))

stargazer(model, out = "./output/Lab3\_Regression/alldata\_age\_weight.html", type = "html",

label = "table:evals-reg",

dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

A table with numbers and symbols

Description automatically generated

All data; hemoglobin ~ age, weight

Спостерігаємо велике значення Constant, що свідчить про необхідність подальшого дослідження впливу інших змінних на гемоглобін.

**Покращення моделей**

Замість побудови та дослідження однієї великої моделі, було вирішено досліджувати 12 окремих моделей.

A diagram of a company

Description automatically generated

Дерево варіантів датасетів (1 датасет = 1 модель)

Моделі відрізняються даними, на яких відбувається побудова.

Перелік наборів даних:

1. sex: male, DRK: Y, SMK: 1
2. sex: male, DRK: Y, SMK: 2
3. sex: male, DRK: Y, SMK: 3
4. sex: male, DRK: N, SMK: 1
5. sex: male, DRK: N, SMK: 2
6. sex: male, DRK: N, SMK: 3
7. sex: female, DRK: Y, SMK: 1
8. sex: female, DRK: Y, SMK: 2
9. sex: female, DRK: Y, SMK: 3
10. sex: female, DRK: N, SMK: 1
11. sex: female, DRK: N, SMK: 2
12. sex: female, DRK: N, SMK: 3

Розбиття початкового датасету:

data <- data %>% filter(age >= 20)

data\_male\_y\_1 <- data %>%

filter(sex == "Male", DRK\_YN == "Y", SMK\_stat\_type\_cd == 1) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_male\_y\_2 <- data %>%

filter(sex == "Male", DRK\_YN == "Y", SMK\_stat\_type\_cd == 2) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_male\_y\_3 <- data %>%

filter(sex == "Male", DRK\_YN == "Y", SMK\_stat\_type\_cd == 3) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_male\_n\_1 <- data %>%

filter(sex == "Male", DRK\_YN == "N", SMK\_stat\_type\_cd == 1) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_male\_n\_2 <- data %>%

filter(sex == "Male", DRK\_YN == "N", SMK\_stat\_type\_cd == 2) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_male\_n\_3 <- data %>%

filter(sex == "Male", DRK\_YN == "N", SMK\_stat\_type\_cd == 3) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_y\_1 <- data %>%

filter(sex == "Female", DRK\_YN == "Y", SMK\_stat\_type\_cd == 1) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_y\_2 <- data %>%

filter(sex == "Female", DRK\_YN == "Y", SMK\_stat\_type\_cd == 2) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_y\_3 <- data %>%

filter(sex == "Female", DRK\_YN == "Y", SMK\_stat\_type\_cd == 3) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_n\_1 <- data %>%

filter(sex == "Female", DRK\_YN == "N", SMK\_stat\_type\_cd == 1) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_n\_2 <- data %>%

filter(sex == "Female", DRK\_YN == "N", SMK\_stat\_type\_cd == 2) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

data\_female\_n\_3 <- data %>%

filter(sex == "Female", DRK\_YN == "N", SMK\_stat\_type\_cd == 3) %>%

select(-sex, -DRK\_YN, -SMK\_stat\_type\_cd)

Для кожного піддатасету побудуємо модель із незалежними змінними: age, weight, height, waistline:

models12 <- list()

for (i in 1:length(data\_12models)) {

model <- lm(hemoglobin ~ age + weight + height + waistline, data = data\_12models[[i]])

model\_hc0 <- coeftest(model, vcov. = vcovHC(model, type = "HC0"))

models12[[i]] = model

print(sprintf("coeftest for model %d", i))

print(model\_hc0)

}

stargazer(models12, out = "./output/Lab3\_Regression/model12\_age\_weight.html", type = "html",

label = "table:evals-reg",

dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

A white sheet with black and white text

Description automatically generated

1. Для деяких наборів даних (8, 9, 10, 11, 12) не всі змінні є попередньо значимими.
2. Наявні великі значення Constant.

Спробуємо зменшити значення Constant шляхом перебору наборів незалежних змінних.

Для кожного піддатасету побудуємо матрицю кореляцій:

A screenshot of a computer generated image

Description automatically generated

На основі матриць кореляцій для кожного піддатасету було обрано незалежні змінні, таким чином, щоб кожна змінна була «попередньо значимою». «Попередня значущість» змінної оцінювалась згідно числової міри кореляції змінної із hemoglobin.

Отримані набори:

1. male\_y\_1: age + height + weight + waistline + DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
2. male\_y\_2: age + height + weight + waistline + DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
3. male\_y\_3: age + height + weight + waistline + DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
4. male\_n\_1: age + height + weight + waistline + hear\_left + hear\_right + DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_AST
5. male\_n\_2: age + height + weight + waistline + hear\_left + hear\_right + DBP + tot\_chole + LDL\_chole + triglyceride + urine\_protein + serum\_creatinine + SGOT\_ALT
6. male\_n\_3: age + height + weight + waistline + DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
7. female\_y\_1: DBP + SBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
8. female\_y\_2: weight + waistline + SBP + DBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
9. female\_y\_3: weight + waistline + SBP + DBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT
10. female\_n\_1: weight + SBP + DBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT + gamma\_GTP
11. female\_n\_2: weight + SBP + DBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT + gamma\_GTP
12. female\_n\_3: weight + waistline + SBP + DBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT

models12 <- list()

models12[[1]] <- lm(hemoglobin ~ age + height + weight + waistline +

DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[1]])

models12[[2]] <- lm(hemoglobin ~ age + height + weight + waistline +

DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[2]])

models12[[3]] <- lm(hemoglobin ~ age + height + weight + waistline +

DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[3]])

models12[[4]] <- lm(hemoglobin ~ age + height + weight + waistline +

hear\_left + hear\_right +

DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_AST,

data = data\_12models[[4]])

models12[[5]] <- lm(hemoglobin ~ age + height + weight + waistline +

hear\_left + hear\_right + DBP + tot\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_ALT,

data = data\_12models[[5]])

models12[[6]] <- lm(hemoglobin ~ age + height + weight + waistline +

DBP + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[6]])

models12[[7]] <- lm(hemoglobin ~ DBP + SBP + BLDS + tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[7]])

models12[[8]] <- lm(hemoglobin ~ weight + waistline + SBP + DBP + BLDS +

tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[8]])

models12[[9]] <- lm(hemoglobin ~ weight + waistline + SBP + DBP + BLDS +

tot\_chole + LDL\_chole + triglyceride + SGOT\_ALT,

data = data\_12models[[9]])

models12[[10]] <- lm(hemoglobin ~ weight + SBP + DBP + BLDS +

tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT + gamma\_GTP,

data = data\_12models[[10]])

models12[[11]] <- lm(hemoglobin ~ weight + SBP + DBP + BLDS +

tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT + gamma\_GTP,

data = data\_12models[[11]])

models12[[12]] <- lm(hemoglobin ~ weight + waistline + SBP + DBP + BLDS +

tot\_chole + LDL\_chole + triglyceride + SGOT\_AST + SGOT\_ALT,

data = data\_12models[[12]])

stargazer(models12, out = "./output/Lab3\_Regression/model12\_manual.html", type = "html",

label = "table:evals-reg",

dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

A screenshot of a computer code

Description automatically generated

1. Спостерігаємо зменшення значень Constant
2. Для моделей (1, 4, 5, 8, 9, 12) деякі незалежні змінні не є значимими

Загалом зменшення Constant не є задовільним, маємо провести подальші дослідження.

Застосуємо AIC для автоматичного підбору незалежних параметрів моделі:

models12 <- list()

for (i in 1:length(data\_12models)) {

min\_model <- lm(hemoglobin ~ 1, data = data\_12models[[i]])

model <- step(min\_model, direction = "forward",

scope = (~ age + height + weight + waistline +

sight\_left + sight\_right + hear\_left + hear\_right +

SBP + DBP + BLDS +

tot\_chole + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine +

SGOT\_AST + SGOT\_ALT + gamma\_GTP))

model\_hc0 <- coeftest(model, vcov. = vcovHC(model, type = "HC0"))

models12[[i]] <- model

print(sprintf("coeftest for model %d", i))

print(model\_hc0)

}

stargazer(models12, out = "./output/Lab3\_Regression/model12\_autobest.html", type = "html",

label = "table:evals-reg",

dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

A screenshot of a computer

Description automatically generated

Constant, майже не відрізняються від значень моделей, параметри яких було обрано вручну.

Очікувалось, що завдяки розбиттю датасету на піддатасети значення Constant буде достатньо малим – адже розбиття гарантує урахування категоріальних змінних (sex, DRK\_YN, SMK\_stat\_type\_cd) у моделі.

Побудуємо 3 моделі без розбиття:

1. AIC, scope = age + height + weight + waistline + sight\_left + sight\_right + hear\_left + hear\_right + SBP + DBP + BLDS + tot\_chole + HDL\_chole + LDL\_chole + triglyceride + urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT + gamma\_GTP
2. AIC, scope = **sex** + age + height + weight + waistline + sight\_left + sight\_right + hear\_left + hear\_right + SBP + DBP + BLDS + tot\_chole + HDL\_chole + LDL\_chole + triglyceride + urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT + gamma\_GTP
3. AIC, scope = **SMK\_Y + SMK\_N** + **DRK\_YN** + **sex** + age + height + weight + waistline + sight\_left + sight\_right + hear\_left + hear\_right + SBP + DBP + BLDS + tot\_chole + HDL\_chole + LDL\_chole + triglyceride + urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT + gamma\_GTP (категоріальна змінна SMK\_stat\_type\_cd (1, 2, 3) замінена двома бінарними змінними SMK\_Y, SMK\_N).

data <- data %>%

mutate(

sex = as.numeric(as.factor(sex)), # male = 2, female = 1

DRK\_YN = as.numeric(as.factor(DRK\_YN)), # 1 = no, 2 = yrs

SMK\_Y = (SMK\_stat\_type\_cd == 3),

SMK\_N = (SMK\_stat\_type\_cd == 1)

) %>%

select(

-SMK\_stat\_type\_cd

)

min\_model <- lm(hemoglobin ~ 1, data = data)

model\_all\_nocat <- step(min\_model, direction = "forward",

scope = (~ age + height + weight + waistline +

sight\_left + sight\_right + hear\_left + hear\_right +

SBP + DBP + BLDS +

tot\_chole + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine +

SGOT\_AST + SGOT\_ALT + gamma\_GTP))

model\_all\_sexcat <- step(min\_model, direction = "forward",

scope = (~ age + height + weight + waistline +

sight\_left + sight\_right + hear\_left + hear\_right +

SBP + DBP + BLDS +

tot\_chole + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine +

SGOT\_AST + SGOT\_ALT + gamma\_GTP +

sex))

model\_all\_allcat <- step(min\_model, direction = "forward",

scope = (~ age + height + weight + waistline +

sight\_left + sight\_right + hear\_left + hear\_right +

SBP + DBP + BLDS +

tot\_chole + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine +

SGOT\_AST + SGOT\_ALT + gamma\_GTP +

sex + DRK\_YN + SMK\_Y + SMK\_N))

stargazer(model\_all\_nocat, model\_all\_sexcat, model\_all\_allcat,

out = "./output/Lab3\_Regression/alldata.html", type = "html",

label = "table:evals-reg",

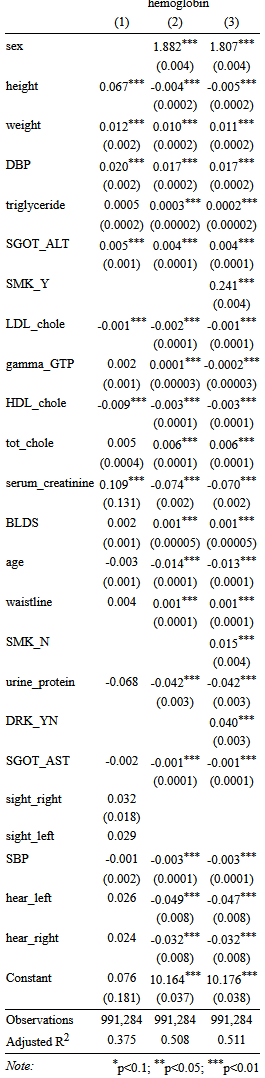
dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

Наявна закономірність: включення категоріальних змінних значно підвищує Constant. Неочікувано, модель (3) включає всі змінні датасету крім власне hemoglobin.

При цьому значення Constant моделі (3) все ще є меншим за значення всіх моделей, побудованих «на розбитті».

Також враховуючи найбільше значення серед всіх побудованих моделей, в якості остаточної моделі оберемо модель (3), побудовану на всьому датасеті.

**Перевірка на мультиколінеарність**

A close-up of a number

Description automatically generated

Наявні суттєві ризики мультиколінеарних змінних. Приберемо змінну tot\_chole:

model\_all\_allcat\_no\_tot\_chole <- lm(hemoglobin ~ age + height + weight + waistline +

hear\_left + hear\_right +

SBP + DBP + BLDS +

HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine +

SGOT\_AST + SGOT\_ALT + gamma\_GTP +

sex + DRK\_YN + SMK\_Y + SMK\_N,

data = data)

stargazer(model\_all\_allcat, model\_all\_allcat\_no\_tot\_chole,

out = "./output/Lab3\_Regression/alldata\_no\_tot\_chole.html", type = "html",

label = "table:evals-reg",

dep.var.labels = c("hemoglobin"),

dep.var.caption = "",

se = list(model\_hc0[, 2]),

no.space = TRUE,

omit.stat = c("rsq", "f", "ser"))

vif(model\_all\_allcat\_no\_tot\_chole)

A close-up of a computer screen

Description automatically generated

A screenshot of a computer code

Description automatically generatedВидалення tot\_chole майже не вплинуло на загальний характер моделі. Ризик мультиколінеарності зменшено.

**Тестування гіпотез**

Для кожної незалежної змінної було протестовано гіпотези:

H0: вплив змінної не є статично значущим

Ha: вплив змінної є статично значущим

Результати тестування:

> hypos <- c('age=0', 'height=0', 'weight=0', 'waistline=0',

+ 'hear\_left=0', 'hear\_right',

+ 'SBP=0', 'DBP=0', 'BLDS=0',

+ 'HDL\_chole=0', 'LDL\_chole=0', 'triglyceride=0',

+ 'urine\_protein=0', 'serum\_creatinine=0',

+ 'SGOT\_AST=0', 'SGOT\_ALT=0', 'gamma\_GTP=0',

+ 'sex=0', 'DRK\_YN=0', 'SMK\_YTRUE=0', 'SMK\_NTRUE=0')

> for (hypo in hypos) {

+ current\_hypos <- hypos[!hypos %in% c(hypo)]

+ print(linearHypothesis(model\_all\_allcat\_no\_tot\_chole,

+ current\_hypos,

+ vcov. = hccm(model\_all\_allcat\_no\_tot\_chole, type='hc1')))

+ }

Linear hypothesis test

Hypothesis:

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 47365 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 21188 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 24676 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 39270 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 53115 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 53098 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 50575 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 47480 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 52426 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 46834 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 52828 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 47363 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 53340 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 28129 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 48290 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 33187 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 46674 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

DRK\_YN = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 4624 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

SMK\_YTRUE = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 42779 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_NTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 40471 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear hypothesis test

Hypothesis:

age = 0

height = 0

weight = 0

waistline = 0

hear\_left = 0

hear\_right = 0

SBP = 0

DBP = 0

BLDS = 0

HDL\_chole = 0

LDL\_chole = 0

triglyceride = 0

urine\_protein = 0

serum\_creatinine = 0

SGOT\_AST = 0

SGOT\_ALT = 0

gamma\_GTP = 0

sex = 0

DRK\_YN = 0

SMK\_YTRUE = 0

Model 1: restricted model

Model 2: hemoglobin ~ age + height + weight + waistline + hear\_left +

hear\_right + SBP + DBP + BLDS + HDL\_chole + LDL\_chole + triglyceride +

urine\_protein + serum\_creatinine + SGOT\_AST + SGOT\_ALT +

gamma\_GTP + sex + DRK\_YN + SMK\_Y + SMK\_N

Note: Coefficient covariance matrix supplied.

Res.Df Df F Pr(>F)

1 991282

2 991262 20 30434 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Для кожної змінної було відкинуто нульову гіпотезу. Тобто вплив кожного параметру є статично значущим, що є очікуваним.

Результати тестування моделі:

coeftest(model\_all\_allcat\_no\_tot\_chole, vcoc. = hccm(model\_all\_allcat\_no\_tot\_chole, type = 'hc0'))

t test of coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.0279e+01 3.7522e-02 273.9331 < 2.2e-16 \*\*\*

age -1.2680e-02 1.0296e-04 -123.1612 < 2.2e-16 \*\*\*

height -4.9152e-03 2.2658e-04 -21.6930 < 2.2e-16 \*\*\*

weight 1.0726e-02 1.6674e-04 64.3289 < 2.2e-16 \*\*\*

waistline 8.7672e-04 1.3233e-04 6.6252 3.469e-11 \*\*\*

hear\_left -4.8816e-02 7.6183e-03 -6.4077 1.478e-10 \*\*\*

hear\_right -3.3866e-02 7.7331e-03 -4.3793 1.190e-05 \*\*\*

SBP -3.2424e-03 1.2075e-04 -26.8525 < 2.2e-16 \*\*\*

DBP 1.7446e-02 1.7174e-04 101.5835 < 2.2e-16 \*\*\*

BLDS 1.3162e-03 4.9097e-05 26.8086 < 2.2e-16 \*\*\*

HDL\_chole 2.4024e-03 8.4843e-05 28.3157 < 2.2e-16 \*\*\*

LDL\_chole 4.3241e-03 3.2273e-05 133.9853 < 2.2e-16 \*\*\*

triglyceride 1.0678e-03 1.2818e-05 83.3061 < 2.2e-16 \*\*\*

urine\_protein -4.2377e-02 2.5763e-03 -16.4484 < 2.2e-16 \*\*\*

serum\_creatinine -6.9573e-02 2.4049e-03 -28.9299 < 2.2e-16 \*\*\*

SGOT\_AST -1.2005e-03 6.3139e-05 -19.0137 < 2.2e-16 \*\*\*

SGOT\_ALT 4.5152e-03 5.9008e-05 76.5180 < 2.2e-16 \*\*\*

gamma\_GTP -1.1522e-04 2.5877e-05 -4.4526 8.486e-06 \*\*\*

sex 1.8034e+00 4.0361e-03 446.8203 < 2.2e-16 \*\*\*

DRK\_YN 4.2028e-02 2.5936e-03 16.2043 < 2.2e-16 \*\*\*

SMK\_YTRUE 2.4268e-01 3.6409e-03 66.6531 < 2.2e-16 \*\*\*

SMK\_NTRUE 1.3525e-02 3.6345e-03 3.7212 0.0001983 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1