

[illegible]

Функція `bartlett.test` використовується для здійснення тесту Бартлетта, який є одним із методів порівняння дисперсій між групами у аналізі даних. Цей тест ставить гіпотезу про рівність дисперсій у різних групах. Якщо значення р-рівня значимості, отримане в результаті тесту, менше за обране значення (зазвичай 0,05 або 0,01), то ми відхиляємо нульову гіпотезу про рівність дисперсій.

В R мова програмування, функція `bartlett.test` використовується для виконання тесту Бартлетта на групах даних для перевірки рівності дисперсій між групами.

В нашому випадку число дорівнює $2.2E-16$ що є менше ніж 0.01 тому ми відхиляємо нульову гіпотезу про рівність дисперсій

Функція ``cohensD`` використовується для обчислення коефіцієнта Коена (Cohen's d), який є одним із показників ефективності різниці між середніми значеннями двох груп у статистичному аналізі.

Коефіцієнт Коена обчислюється як різниця між середніми значеннями двох груп, поділена на загальний стандартний відхил. Він дозволяє оцінити розмір ефекту зміни між групами у стандартних відхиленнях.

У більшості випадків, значення коефіцієнта Коена менше 0.2 вказує на дуже малий ефект, 0.2-0.5 - на малий ефект, 0.5-0.8 - на середній ефект, а більше 0.8 - на великий ефект.

В нашому випадку значення 0.76 тому маємо великий ефект

```
> lm(formula ~ data$bedroom_count)
```

Call:

```
lm(formula = formula ~ data$bedroom_count)
```

Coefficients:

```
(Intercept) data$bedroom_count  
6.276e+14    -2.946e+13
```

Це результат лінійної регресії (моделі), який вказує на оцінки коефіцієнтів для рівняння регресії, що моделює зв'язок між двома змінними: залежною змінною (`formula`) і незалежною змінною (`data$bedroom_count`).

У вищенаведеному виразі `lm` (`linear model`) показує, що наше рівняння регресії виглядає приблизно так:

```
\[ \text{formula} = 6.276 \times 10^{14} - 2.946 \times 10^{13} \times  
\text{bedroom_count} \]
```

Це означає, що для кожного одиниці зміни у `bedroom_count`, `formula` зменшується на приблизно 2.946×10^{13} , ураховуючи інші змінні. Коефіцієнт перед першим членом є значенням перетину з віссю у, коли всі інші змінні рівні нулю.

```
> summary(lm)$coefficients[, 2]  
  (Intercept) data$bedroom_count  
  2.624523e+13  5.868340e+12
```

Це результати коефіцієнтів змінної нахилу (slope) та константи (intercept), отримані з використанням функції `summary(lm)$coefficients[, 2]`.

Згідно з цими результатами:

- Значення коефіцієнта перед `(Intercept)` (константа) приблизно дорівнює 2.624523×10^{13} .
- Значення коефіцієнта перед `data$bedroom_count` (змінна нахилу) приблизно дорівнює 5.868340×10^{12} .

Ці значення вказують на те, як змінюється наша залежна змінна (`formula`) при зміні на одиницю нашої незалежної змінної (`data$bedroom_count`), ураховуючи інші змінні в моделі.

```
X <- cbind(1,data$bedroom_count)
var_betaHat <- anova(lm)[[3]][2] * solve(t(X) %*% X)
sqrt(diag(var_betaHat))
```

Цей код використовується для обчислення стандартних помилок коефіцієнтів регресії у лінійній моделі. Ось пояснення кожного кроку:

`X <- cbind(1,data$bedroom_count)`: Тут створюється матриця `X`, яка містить перший стовпчик з одиницями (це для константи) і другий стовпчик зі значеннями `data$bedroom_count` (це для незалежної змінної в моделі).

`var_betaHat <- anova(lm)[[3]][2] * solve(t(X) %*% X)`:

Вираховується дисперсія коефіцієнта перед `data$bedroom_count` у регресійній моделі. `anova(lm)[[3]][2]` повертає квадрат стандартної помилки для `data$bedroom_count` з результатів аналізу дисперсії (`anova(lm)`). `solve(t(X) %*% X)` обчислює обернену матрицю, яка використовується для обчислення коваріаційної матриці для оцінок коефіцієнтів.

`sqrt(diag(var_betaHat))`: Обчислюється квадратний корінь з діагональних елементів матриці `var_betaHat`, що представляють стандартні помилки кожного коефіцієнта.

Отже, результат цієї операції дає нам стандартні помилки кожного коефіцієнта у нашій моделі лінійної регресії.