

BÀI TẬP THỰC HÀNH 2

THỐNG KÊ MÔ TẢ

1. Đọc file dữ liệu

- Sử dụng file dữ liệu **igfdata**

Ngoài các chỉ số liên quan đến giới tính, độ tuổi, trọng lượng và chiều cao, chúng tôi đo lường các hormone liên quan đến tình trạng tăng trưởng như **igfi**, **igfbp3**, **als**, và các markers liên quan đến sự chuyển hóa của xương **pinp**, **ictp** và **pinp**

```
> setwd("E:/Khoa CNTT/Lecturers/Data analysis/Data")  
> getwd()
```

- Đọc vào một dataframe

```
> igfdata <- read.table("igf.txt", header=TRUE,  
na.strings=".")  
> attach(igfdata)  
> names(igfdata)  
> igfdata
```

2. Tần số (xác suất)

2.1. Thao tác cơ bản

- Hiển thị tần số của biến số phân loại như: sex, ethnicity.
- Hiển thị bảng thống kê 2 chiều cho 2 biến: sex, ethnicity.
- Hiển thị số lượng và tỉ lệ người là nam (Male) trong dataframe đang xét

2.2. Kiểm định tần số

Trong nghiên cứu trên, chúng ta thấy có 69 nữ và 31 nam. Như vậy tỉ lệ nữ là 0.69 (hay 69%). Để kiểm định xem tỉ lệ này có thật sự khác với tỉ lệ 0.5 hay không?