

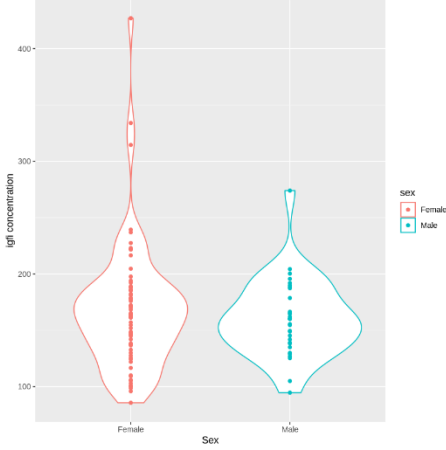
TRƯỜNG ĐẠI HỌC VĂN LANG
Khoa Công nghệ ứng dụng

ĐỀ THI, ĐÁP ÁN/RUBRIC VÀ THANG ĐIỂM
THI KẾT THÚC HỌC PHẦN
Học kỳ 2, năm học 2023-2024

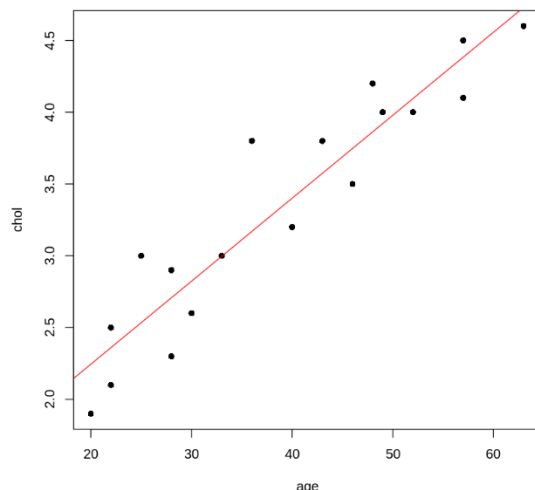
I. Thông tin chung

Tên học phần:	Thống kê sinh học		
Mã học phần:	71GENE40222	Số tin chỉ:	2
Mã nhóm lớp học phần:	232_71GENE40222_01		
Hình thức thi: Tự luận	Thời gian làm bài:	90	phút
Thí sinh được tham khảo tài liệu:	☒ Có	☐ Không	

ĐÁP ÁP VÀ THANG ĐIỂM

Phần câu hỏi	Nội dung đáp án	Thang điểm	Ghi chú
I. Tự luận			
Câu 1	- Có 100 bệnh nhân và 12 biến - Biến nhị phân (sex), biến phân loại (ethnicity), biến rời rạc (age), biến liên tục (height, weight, igfi, igfbp3, als, pinp, ictp, p3np) - Trung bình, trung vị độ tuổi: 38.83, 38 - Trung bình độ tuổi của nam (19.59) và nữ (18.23) - Nồng độ trung bình của hormone igfi ở nam (167.97) và nữ (160.29)	2.5	
Câu 2	- Biểu đồ về phân bố nồng độ hormone igfi ở nam và nữ. 	2.5	

	- Kết quả kiểm định t-test cho thấy không có sự khác biệt nào về hormone igfi giữa nam và nữ Welch Two Sample t-test data: igfi by sex $t = 0.84117$, $df = 88.329$, $p\text{-value} = 0.4025$ alternative hypothesis: true difference in means between group Female and group Male is not equal to 0 95 percent confidence interval: -10.46855 25.83627 sample estimates: mean in group Female mean in group Male 167.9741 160.2903 - Xác định số mẫu cần cho nghiên cứu cho dựa trên trung bình chỉ số igfi giữa nam và nữ Two-sample t test power calculation $n = 801.8676$ $\delta = 7$ $sd = 50$ $\text{sig.level} = 0.05$ $\text{power} = 0.8$ $\text{alternative} = \text{two.sided}$ NOTE: n is number in *each* group Kết luận: cần thu thập 801 bệnh nhân ở mỗi nhóm nam và nữ để xác định mức độ khác biệt về chỉ số igfi với mức ảnh hưởng 0.8 và mức ý nghĩa <0.05		
Câu 3	- Kết quả kiểm định Student t-test cho biến theo cặp Paired t-test data: before and after $t = 2.7924$, $df = 9$, $p\text{-value} = 0.02097$ alternative hypothesis: true mean difference is not equal to 0 95 percent confidence interval: 1.993901 19.006099 sample estimates: mean difference 10.5 - Kết quả trên cho thấy sau khi điều trị áp suất máu giảm 10.5 mmHg, và khoảng tin cậy 95% là từ 2.0 mmHg đến 19 mmHg, với trị số $p = 0.0209$. Như vậy, chúng ta có bằng chứng để phát biểu rằng mức độ giảm huyết áp có ý nghĩa thống kê.	2.5	
Câu 4	# Vẽ biểu đồ tán xạ (scatter plot) giữa độ tuổi và cholesterol $\text{plot}(\text{chol} \sim \text{age}, \text{pch}=16)$ # Vẽ đường hồi quy giữa 2 biến $\text{abline}(\text{lm}(\text{chol} \sim \text{age}), \text{col}=\text{"red"})$	2.5	



Xác định hệ số tương quan Pearson
cor(age, chol)

0.93

- Có sự tương quan mạnh giữa độ tuổi và cholesterol

- Kết quả kiểm định giả thiết hệ số tương quan bằng 0 (tức hai biến x và y không có liên hệ) cho thấy sự tương quan giữa độ tuổi và cholesterol có ý nghĩa về mặt thống kê với $P = 1.058e-08$ và ngưỡng tương quan từ 0.83-0.97.

Xây dựng mô hình hồi quy tuyến tính giữa độ tuổi và cholesterol

Call: lm(formula = chol ~ age)

Residuals: Min 1Q Median 3Q Max -0.40729 -0.24133 -0.04522 0.17939 0.63040 Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.089218 0.221466 4.918 0.000154
*** age 0.057788 0.005399 10.704 1.06e-08 ***
--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 0.3027 on 16 degrees of freedom Multiple R-squared: 0.8775, Adjusted R-squared: 0.8698 F-statistic: 114.6 on 1 and 16 DF, p-value: 1.058e-08

- Mô hình hồi quy tuyến tính giữa độ tuổi và cholesterol là: $y = 0.05x + 1.08$ ($a=0.05$, $b=1.08$)

Điểm tổng

10.0