108年台北醫學大學生物統計學分班第七次作業

Bo-Jiang Lin

2019-9-16

試利用心血管疾病資料（CVD ALL）分析沒有心血管疾病成人之腰圍與收縮壓的關係。（顯著水準設定為0.05）

1. 請問腰圍與收縮壓的皮爾生相關係數為何？兩者是否存在顯著的線性關係？

cor.test(HW7$Compression, HW7$Waistline, use = "complete.obs", method = "pearson")

##   
## Pearson's product-moment correlation  
##   
## data: HW7$Compression and HW7$Waistline  
## t = 117.66, df = 62381, p-value < 2.2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.4197215 0.4325658  
## sample estimates:  
## cor   
## 0.4261651

# contain use = "complete.obs" when missing value exist

1. 請問腰圍與收縮壓的斯皮爾曼等級相關為何？兩者是否存在顯著的等級相關？

cor.test(HW7$Compression, HW7$Waistline, use = "complete.obs", method = "spearman")

## Warning in cor.test.default(HW7$Compression, HW7$Waistline, use =  
## "complete.obs", : Cannot compute exact p-value with ties

##   
## Spearman's rank correlation rho  
##   
## data: HW7$Compression and HW7$Waistline  
## S = 2.1942e+13, p-value < 2.2e-16  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.4577024

1. 試利用簡單線性迴歸模型建立腰圍預測收縮壓之模型

Waistline = HW7$Waistline  
regresion <- lm(HW7$Compression ~ Waistline, na.action = na.omit)  
summary(regresion)

##   
## Call:  
## lm(formula = HW7$Compression ~ Waistline, na.action = na.omit)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -112.092 -12.824 -2.279 10.344 147.215   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 58.140635 0.558520 104.1 <2e-16 \*\*\*  
## Waistline 0.831113 0.007064 117.7 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.82 on 62381 degrees of freedom  
## (2106 observations deleted due to missingness)  
## Multiple R-squared: 0.1816, Adjusted R-squared: 0.1816   
## F-statistic: 1.384e+04 on 1 and 62381 DF, p-value: < 2.2e-16

1. 請問此模型為何？

cat("y =", as.numeric(regresion$coeff[2]), "x+", as.numeric(regresion$coeff[1]))

## y = 0.831113 x+ 58.14064

# using cat() to print a sentence with multiple variables

2. 腰圍是否與收縮壓有顯著相關？此模型腰圍解釋了收縮壓變異的多少百分比？

cat("Adjusted R-square is ", summary(regresion)$adj.r.squared,",腰圍和收縮壓為顯著性相關，然腰圍僅仍解釋", as.numeric(summary(regresion)$adj.r.squared)\*100,"%的收縮壓變異")

## Adjusted R-square is 0.1816036 , 腰圍和收縮壓為顯著性相關，然腰圍僅仍解釋 18.16036 %%的收縮壓變異

3. 若有一人腰圍為100公分，請預測此人平均而言收壓縮何？

newdata = data.frame("Waistline" = 100)  
predict(regresion, newdata , interval = "prediction", na.action = na.exclude)

## fit lwr upr  
## 1 141.2519 104.3624 178.1415