

第十次作业

1 第一题

收集某学校20名一年级女大学生肺活量(L)、胸围(cm)和数据，假设两变量均近似服从正态分布，问二者间有无线性关联。

依据题意，做出检验假设

H_0 : 总体相关系数 $\rho = 0$, 即认为肺活量和胸围之间线性无关

H_1 : 总体相关系数 $\rho \neq 0$, 即认为肺活量和胸围之间线性相关

$\alpha = 0.05$

(1)

从而进行检验：

```
d ← matrix(c(
  2.99, 51.3, 73.6,
  3.11, 48.9, 83.9,
  1.91, 42.8, 78.3,
  2.63, 55.0, 77.1,
  2.86, 45.3, 81.7,
  1.91, 45.3, 74.8,
  2.98, 51.4, 73.7,
  3.28, 53.8, 79.4,
  2.52, 49.0, 72.6,
  3.27, 53.9, 79.5,
  3.10, 48.8, 83.8,
  3.28, 52.6, 88.4,
  1.92, 42.7, 78.2,
  3.27, 52.5, 88.3,
  2.64, 55.1, 77.2,
  2.85, 45.2, 81.6,
  3.16, 51.4, 78.3,
  2.51, 48.7, 72.5,
  3.15, 51.3, 78.2,
  1.92, 45.2, 74.7
), 20, 3, byrow = TRUE)
# vital_capacity
vc ← d[, 1]
# weight
wg ← d[, 2]
# chest_girth
cg ← d[, 3]

plot(vc, cg,
```

```

main = "vital_capacity vs chest_girth",
xlab = "vital_capacity",
ylab = "chest_girth"
)

cor.test(vc, cg)

```

得到输出结果 $r = 0.5155538$ ，带入公式 $t = \frac{r-0}{s_r} \sim t(n-2)$ ，得到结果 $t = 2.5527, p\text{-value} = 0.01999 < 0.05 = \alpha$ ，从而按照 $\alpha = 0.05$ 标准，拒绝 H_0 ，接受 H_1 ，也即认为总体相关系数 $\rho \neq 0$ ，两者之间具有线性相关关系。

2 第二题

以女大学生肺活量(L)为因变量，胸围(cm)为自变量，作简单线性回归分析。

两变量具有线性相关关系的检验已经在第一题得到确认。

```

lm(vc ~ cg)
summary(lm(vc ~ cg))

```

根据输出结果，得到拟合直线 $vc = 0.05468cg - 1.54535$ ，回归系数 $b = 0.05468$ ，即平均而言，胸围每增加1cm，肺活量增加0.05468L。

对回归系数 b 做 t 检验

建立检验假设

$$\begin{aligned}
 H_0 : \beta &= 0, \text{ 即方程不成立} \\
 H_1 : \beta &\neq 0, \text{ 即方程成立} \\
 \alpha &= 0.05
 \end{aligned}
 \tag{2}$$

同时，根据公式 $t = \frac{b-\beta}{s_b} \sim t(n-2)$ 对回归系数 b 做 t 检验，得到 $t = 2.553, Pr(> t) = 0.020 < \alpha = 0.05$ ，从而拒绝 H_0 ，接受 H_1 ，也即认为方程成立。

3 第三题

计算胸围为79(cm)的一年级女大学生，平均肺活量估计值是多少？平均肺活量的95%置信区间是多少？肺活量在什么范围内？

```

predict(lm(vc ~ cg), data.frame(cg = c(79)), interval = c("confidence"))
predict(lm(vc ~ cg), data.frame(cg = c(79)), interval = c("prediction"))

```

从而得到结果，平均肺活量估计值为 $2.774483L$ ，平均肺活量的95%置信区间 $2.567745L \sim 2.981221L$ 。个体肺活量的95%置信区间为 $1.828035L \sim 3.720931L$