

第四次作业

1 第一题

已知一般健康成年女性血红蛋白的均数为124.7g/L，某医生在某山区随机抽取了20例健康成年女性，测得她们血红蛋白如下，问：该山区健康成年女性血红蛋白均数是否与一般健康成年女性不同？

该山区20例健康成年女性血红蛋白（g/L）为：110 100 120 115 130 120 140 110 125 100 115 125 95 115 110 125 110 115 115 105

1.1 单样本t检验

1.1.1 建立检验假设

$$\begin{aligned} H_0: \mu &= \mu_0, \text{即假设该山区健康成年女性血红蛋白均数与一般健康成年女性相同} \\ H_1: \mu &\neq \mu_0, \text{即假设该山区健康成年女性血红蛋白均数与一般健康成年女性不同} \end{aligned} \quad (1)$$

1.1.2 确定检验水准

设立显著性水准 $\alpha = 0.05$ ，选择单样本t检验

1.1.3 确定p值，做出推断结论

```
hemo <- read.csv("Homework4/Data/Part1.csv")$Hemoglobin
summary(hemo)

t.test(hemo, mu = 124.7) # t = -3.9863, df = 19, p-value = 0.0007905
```

得到结果 $t = -3.9863, df = 19, p\text{-value} = 0.0007905$

1.1.4 总结

依题意，建立检验假设

$$\begin{aligned} H_0: \mu &= \mu_0 \\ H_1: \mu &\neq \mu_0 \end{aligned} \quad (2)$$

其中无效假设 H_0 认为样本均数与总体均数相等，备择假设 H_1 认为样本均数与总体均数不相等。

确立检验水准，设立显著性水准 $\alpha = 0.05$ ，选择单样本t检验，带入公式

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}} \quad v = n - 1 = 19 \quad (3)$$

得 $t = -3.9863, df = 19, p\text{-value} = 0.0007905 < 0.05 = \alpha$ ，拒绝 H_0 ，接受 H_1 ，认为样本均数与总体均数不相等。

得出结论：山区健康成年女性血红蛋白均数与一般健康成年女性不同。

2 第二题

某医院用某种中草药治疗高血压病人10名，治疗前后舒张压的变化如下表，问该中草药对于降低舒张压有无作用？

2.1 配对设计的t检验

2.1.1 建立检验假设

$$\begin{aligned}H_0: \mu_d &= 0 \\H_1: \mu_d &\neq 0\end{aligned}\quad (4)$$

2.1.2 确定检验水准

设立显著性水准 $\alpha = 0.05$ ，选择单样本t检验

2.1.3 确定p值

```
treat <- read.csv("Homework4/Data/Part2.csv")
before <- treat$before
after <- treat$after

diff <- after - before
summary(diff)

t.test(diff, mu = 0) # t = -3.7154, df = 9, p-value = 0.004804
# 需要报告95%CI
```

得到结果 $t = -3.7154, df = 9, p\text{-value} = 0.004804$

2.1.4 总结

依题意，建立检验假设

$$\begin{aligned}H_0: \mu_d &= 0 \\H_1: \mu_d &\neq 0\end{aligned}\quad (5)$$

其中无效假设 H_0 认为两总体均数相等，备择假设 H_1 认为两总体均数不相等。

确立检验水准，设立显著性水准 $\alpha = 0.05$ ，选择单样本t检验，带入公式

$$\begin{aligned}d &= X_{after} - X_{before} \quad n = 11 \\t &= \frac{\bar{d} - 0}{s_d / \sqrt{n}} \quad v = n - 1 = 10\end{aligned}\quad (6)$$

得 $t = -3.7154, df = 9, p\text{-value} = 0.004804 < 0.05 = \alpha$ ，拒绝 H_0 ，接受 H_1 ，认为两总体均数不相等。

得出结论：该中草药对于降低舒张压有作用。

3 第三题

某医生从OC服用者中随机抽取8名35~39岁妇女，另从未服用OC者中随机抽取19名35~39岁妇女，测量这两组妇女的血压，得到如下表资料，问服用OC与否是否对35~39岁妇女的血压有影响？

3.1 两个独立样本的t检验

3.1.1 建立检验假设

$$\begin{aligned}H_0: \mu_1 &= \mu_2 \\H_1: \mu_1 &\neq \mu_2\end{aligned}\quad (7)$$

其中无效假设 H_0 认为两独立样本总体均数相等，备择假设 H_1 认为两独立样本总体均数不相等。

3.1.2 确定检验水准

设立显著性水准 $\alpha = 0.05$ ，选择两独立样本的t检验

3.1.3 检验方差齐性

3.1.3.1 建立检验假设

$$\begin{aligned}H_0: \sigma_1^2 &= \sigma_2^2 \\H_1: \sigma_1^2 &\neq \sigma_2^2\end{aligned}\quad (8)$$

3.1.3.2 确定检验水准

设立显著性水准 $\alpha = 0.10$ ，选择F检验

3.1.3.3 使用F分布校验

$$F = \frac{S_1^2}{S_2^2} \quad v_1 = n_1 - 1, v_2 = n_2 - 1 \quad (9)$$

```
medicine <- read.csv("Homework4/Data/Part3.csv")
used <- na.omit(medicine$Used)
not_used <- medicine$NotUsed

var.test(used, not_used)
```

得到结果 `F = 1.2511, num df = 7, denom df = 18, p-value = 0.655`

3.1.3.4 得出结论

因为 $p = 0.655 > 0.1 = \alpha$ ，接受 H_0 ，认为两独立样本满足方差齐性。

3.1.4 进行两独立样本的t检验，确定p值

```
t.test(used, not_used, var.equal = TRUE)
```

得到结果 `t = 1.4769, df = 25, p-value = 0.1522`

3.1.5 总结

依题意，建立检验假设

$$\begin{aligned}H_0: \sigma_1^2 &= \sigma_2^2 \\H_1: \sigma_1^2 &\neq \sigma_2^2\end{aligned}\tag{10}$$

其中无效假设 H_0 认为两独立样本总体均数相等，备择假设 H_1 认为两独立样本总体均数不相等。

确立检验水准，设立显著性水准 $\alpha = 0.05$ ，选择双样本t检验，先进行方差齐性检验，知两独立样本满足方差齐性，然后带入公式

$$\begin{aligned}t &= \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} \\s_{\bar{X}_1 - \bar{X}_2} &= \sqrt{s_c^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \\s_c^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\end{aligned}\tag{11}$$

得 $t = 1.4769, df = 25, p - value = 0.1522 > 0.05 = \alpha$ ，接受 H_0 ，拒绝 H_1 ，认为两总体均数相等。

得出结论：服用OC与否是否对35~39岁妇女的血压没有影响。

4 附录

以下为全部作业代码

```
# Part 1
hemo <- read.csv("Homework4/Data/Part1.csv")$Hemoglobin
summary(hemo)

t.test(hemo, mu = 124.7) # t = -3.9863, df = 19, p-value = 0.0007905

# Part 2
treat <- read.csv("Homework4/Data/Part2.csv")
before <- treat$before
after <- treat$after

diff <- after - before
summary(diff)

t.test(diff, mu = 0) # t = -3.7154, df = 9, p-value = 0.004804
```

```
# Part 3
medicine <- read.csv("Homework4/Data/Part3.csv")
used <- na.omit(medicine$Used)
not_used <- medicine$NotUsed
summary(used)

var.test(used, not_used)
t.test(used, not_used, var.equal = TRUE)
```