

量化研究方法模拟试题

考试说明：本试卷共三道大题。请根据各题提供的背景和数据集，将其保存为 CSV 文件后，使用 R 语言进行统计分析，并根据输出结果回答问题。

第一题：独立样本 t 检验 (Independent Samples t-test)

1.1 题目背景

研究方言 A 母语者在发低元音 /a/ 时，男性和女性的基频 F0 是否存在显著差异。随机抽取 40 名受试者（男女各 20 名）。

1.2 数据集 (请保存为 task1_f0_data.csv)

ID	Gender	F0	ID	Gender	F0
1	Male	120	21	Female	210
2	Male	115	22	Female	225
3	Male	128	23	Female	198
4	Male	110	24	Female	205
5	Male	132	25	Female	240
6	Male	105	26	Female	215
7	Male	118	27	Female	220
8	Male	122	28	Female	212
9	Male	125	29	Female	208
10	Male	119	30	Female	230
11	Male	112	31	Female	218
12	Male	126	32	Female	202
13	Male	130	33	Female	224
14	Male	117	34	Female	211
15	Male	121	35	Female	219
16	Male	124	36	Female	227
17	Male	113	37	Female	206
18	Male	116	38	Female	214
19	Male	129	39	Female	221
20	Male	123	40	Female	228

第二题：单因素方差分析 (One-way ANOVA)

2.1 题目背景

对比三种词汇学习策略（A: 机械记忆, B: 语境推断, C: 游戏化）对学生成绩的影响（满分 100）。

2.2 数据集 (请保存为 task2_strategy_data.csv)

Group	Score	Group	Score	Group	Score
A	65	B	78	C	88
A	68	B	82	C	92
A	70	B	75	C	85
A	62	B	80	C	90
A	71	B	85	C	86
A	64	B	77	C	89
A	66	B	79	C	91
A	69	B	81	C	87
A	63	B	83	C	84
A	67	B	76	C	93
A	72	B	84	C	89
A	60	B	80	C	90
A	65	B	78	C	88
A	68	B	82	C	86
A	70	B	79	C	95

第三题：多元线性回归分析 (Multiple Regression Analysis)

3.1 题目背景

预测学生情绪评分 (Mood, 0-100) 如何受学习时长 (Study)、睡眠时长 (Sleep)、社交频率 (Social) 和焦虑水平 (Anxiety) 的影响。

3.2 完整数据集 (请保存为 task3_mood_data.csv)

ID	Study	Sleep	Social	Anxiety	Mood	ID	Study	Sleep	Social	Anxiety	Mood
1	15	7.0	2	20	75	26	14	7.2	2	22	74
2	25	6.0	1	35	60	27	28	5.5	1	40	55
3	10	8.5	4	15	88	28	12	8.0	3	18	82
4	30	5.0	0	45	45	29	22	6.5	2	30	68
5	18	7.5	3	22	80	30	16	7.8	4	15	85
6	20	6.8	2	28	70	31	19	7.0	2	25	72
7	12	8.2	5	10	92	32	11	8.5	5	12	90
8	22	6.5	1	32	65	33	24	6.2	1	38	58
9	17	7.2	3	20	78	34	15	7.5	3	20	79
10	28	5.8	0	42	52	35	26	6.0	0	44	50
11	14	7.8	4	18	83	36	13	8.2	4	14	87
12	26	6.2	2	30	66	37	21	6.8	2	26	71
13	19	7.0	3	25	74	38	18	7.4	3	21	77
14	16	7.4	2	23	76	39	17	7.6	2	24	75
15	23	6.4	1	35	62	40	25	6.3	1	36	61
16	11	8.0	5	12	89	41	12	8.4	5	11	93
17	29	5.2	0	48	42	42	30	5.4	0	46	46
18	21	7.2	2	26	73	43	19	7.1	3	22	78
19	13	8.1	4	15	86	44	15	7.7	4	16	84
20	24	6.0	1	37	59	45	27	5.9	1	41	54
21	18	7.3	3	21	79	46	20	6.9	2	27	70
22	20	6.6	2	29	67	47	14	7.9	3	19	81
23	10	8.6	6	8	95	48	23	6.5	1	34	63
24	27	5.6	0	43	48	49	11	8.7	6	9	96
25	15	7.5	3	19	80	50	29	5.1	0	49	41

附录：R 语言参考答案代码

```
# =====  
# 环境准备：请确保将数据文件放在 R 的工作目录下  
# =====  
  
# 如果需要读取 .xlsx 文件，请取消下面两行的注释：  
# install.packages("readxl")  
# library(readxl)  
  
# =====  
# 题目 1: 独立样本 t 检验  
# =====  
  
# 1. 导入数据 (假设为 CSV)  
# data_t <- read_excel("task1_f0_data.xlsx") # 如果是 Excel  
data_t <- read.csv("task1_f0_data.csv")  
  
# 2. 描述性统计  
library(psych)  
describeBy(data_t$F0, data_t$Gender)  
  
# 3. 正态性检验 (Shapiro-Wilk)  
by(data_t$F0, data_t$Gender, shapiro.test)  
  
# 4. 执行 t 检验 (假设方差齐性)  
t_test_results <- t.test(F0 ~ Gender, data = data_t, var.equal = TRUE)  
print(t_test_results)  
  
# =====  
# 题目 2: 单因素 ANOVA  
# =====  
  
# 1. 导入数据  
data_anova <- read.csv("task2_strategy_data.csv")
```

```
data_anova$Group <- as.factor(data_anova$Group)
```

```
# 2. 方差齐性检验
```

```
library(car)
```

```
leveneTest(Score ~ Group, data = data_anova)
```

```
# 3. ANOVA 分析
```

```
anova_model <- aov(Score ~ Group, data = data_anova)
```

```
summary(anova_model)
```

```
# 4. 事后检验 (如果  $p < 0.05$ )
```

```
TukeyHSD(anova_model)
```

```
# 5. 可视化
```

```
boxplot(Score ~ Group, data = data_anova, col = "lightblue",  
        main = "Scores by Learning Strategy", xlab = "Strategy", ylab = "Score")
```

```
# =====
```

```
# 题目 3: 多元线性回归分析
```

```
# =====
```

```
# 1. 导入数据
```

```
data_reg <- read.csv("task3_mood_data.csv")
```

```
# 2. 建立回归模型
```

```
# 因变量: Mood; 自变量: Study, Sleep, Social, Anxiety
```

```
mood_model <- lm(Mood ~ Study + Sleep + Social + Anxiety, data = data_reg)
```

```
# 3. 查看详细结果 (包含 R-squared 和 p-value)
```

```
summary(mood_model)
```

```
# 4. 检查多重共线性 (VIF)
```

```
library(car)
```

```
vif(mood_model)
```