

レポート提出票

科目名: 応用情報工学演習

実験課題名: 寒水研究室 第6-7回

実施日: 2025年 10月 22日

共同実験者

4623044 呉 子吟

演習課題

課題 1 連続量のデータ「age」,「sbp_b」,「dbp_b」,「sbp_a」,「dbp_a」の基本統計量（データ数, 平均, 標準偏差, 中央値, 最小値, 最大値）を群ごとに計算して, 結果をみやすい表にして整理して. さらに, 各変数のヒストグラムの形を確認してください（レポートには不要です）.

2 値変数のデータ「sex」の各水準の度数と割合を群ごとに計算してください.

課題 2 連続量のデータについて, 散布図や相関係数をもとに, 変数間の関係を調べてください. とくに, 2 か月後の収縮期血圧と関連する変数があるかどうかを確認してください.

課題 3 横軸を群, 縦軸を 2 か月後の収縮期血圧とした散布図を作成してください. 2 つの群の収縮期血圧の分布が見やすいようにグラフの表示を工夫してください.

課題 4 課題 1 の結果から, 開始時点の収縮期血圧, 年齢, 性別が, 2 つの群間でバランスがとれているかを確認してください.

課題 1 の結果

付録 1 に示した Listing1 の R コードを用いて、薬群とプラセボ群の年齢および血圧に関する基本統計量を計算し、その結果、表 1 に示されている。

表 1: 群別の基本統計量

グループ	データ数	平均	標準偏差	中央値	最小値	最大値
Age						
Drug group	5	45.6	4.62	43	42	52
Placebo group	5	58.4	6.77	55	52	69
SBP(b)						
Drug group	5	151.6	2.19	151	149	155
Placebo group	5	165.8	4.76	167	161	172
DBP(b)						
Drug group	5	96.0	3.39	96	91	100
Placebo group	5	97.6	8.73	98	88	111
SBP(a)						
Drug group	5	151.6	2.19	151	149	155
Placebo group	5	166.8	12.15	164	154	185
DBP(a)						
Drug group	5	96.0	3.39	96	91	100
Placebo group	5	102.2	7.19	103	93	112

表 2: 群別の性別分布

Group	性別	度数	相対度数 (%)
Drug group	F	4	80
	M	1	20
Placebo group	F	3	60
	M	2	40

2 値変数である sex（性別）について、各群における度数および割合を算出した（表 2）。薬群では女性が 80%，男性が 20%，プラセボ群では女性が 60%，男性が 40%であったことがわかる。両群とも女性の方が多い傾向があり、性別の偏りがあると考えている。このような偏りは実験結果の解釈に影響する可能性があり、実際の臨床試験においては、場合によって性別構成のバランスが崩れないように被験者を割り付けることが望ましい。

課題 2 の結果

Pairwise Scatter Plots

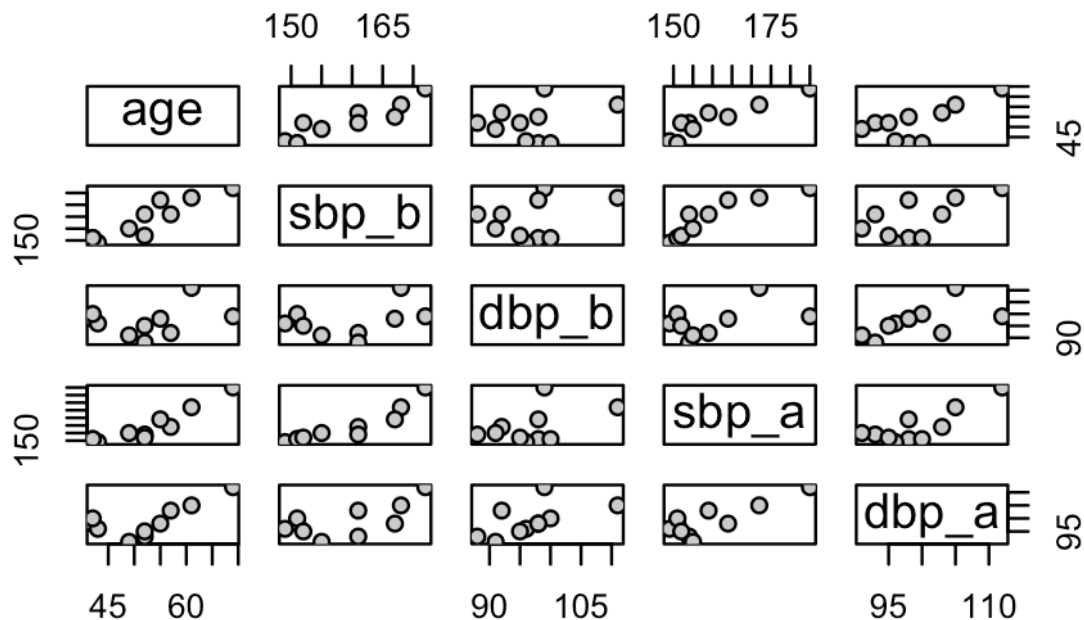


図 1: 2 変数の散布図

表 3: 2 変数ペアの相関行列

Variable	age	sbp_b	dbp_b	sbp_a	dbp_a
age	1.000	0.921	0.274	0.932	0.689
sbp_b	0.921	1.000	0.315	0.912	0.651
dbp_b	0.274	0.315	1.000	0.461	0.578
sbp_a	0.932	0.912	0.461	1.000	0.825
dbp_a	0.689	0.651	0.578	0.825	1.000

課題 2 で連続量のデータ age, sbp_b, dbp_b, sbp_a, dbp_a の関係性を確認するために、2 変数ごとのペア散布図を作成し、図 1 のように 2 変数間の関係を示している。また、R 言語の `cor` 関数を用いて、表 3 のような 2 変数ごと相関係数を計算した。

特に、図 1 の散布図と表 3 より、その結果、2 か月後の収縮期血圧 sbp_a は他の変数と正の相関があることが確認された。特に、介入前の収縮期血圧 (sbp_b)、年齢 (age)、との間では $r > 0.9$ と非常に高い相関が確認され、初期の血圧や年齢が高いほど、2 か月後の血圧も高く維持される傾向があることが示唆された。

さらに、2 か月後の拡張期血圧 (dbp_a($r = 0.82$) の相関関係も強いが、実験前 dbp_b) についても中低程度の相関 ($r = 0.46$) だけ認められることがわかる。

これらの結果から、2 か月後の収縮期血圧 (sbp_a) は介入前の血圧と年齢の影響を受けており、被験者の基礎的な身体状況と年齢が介入効果の評価に影響を及ぼす可能性があると考えられる。

課題 3 の結果

Systolic Blood Pressure by Group (After 2 months)

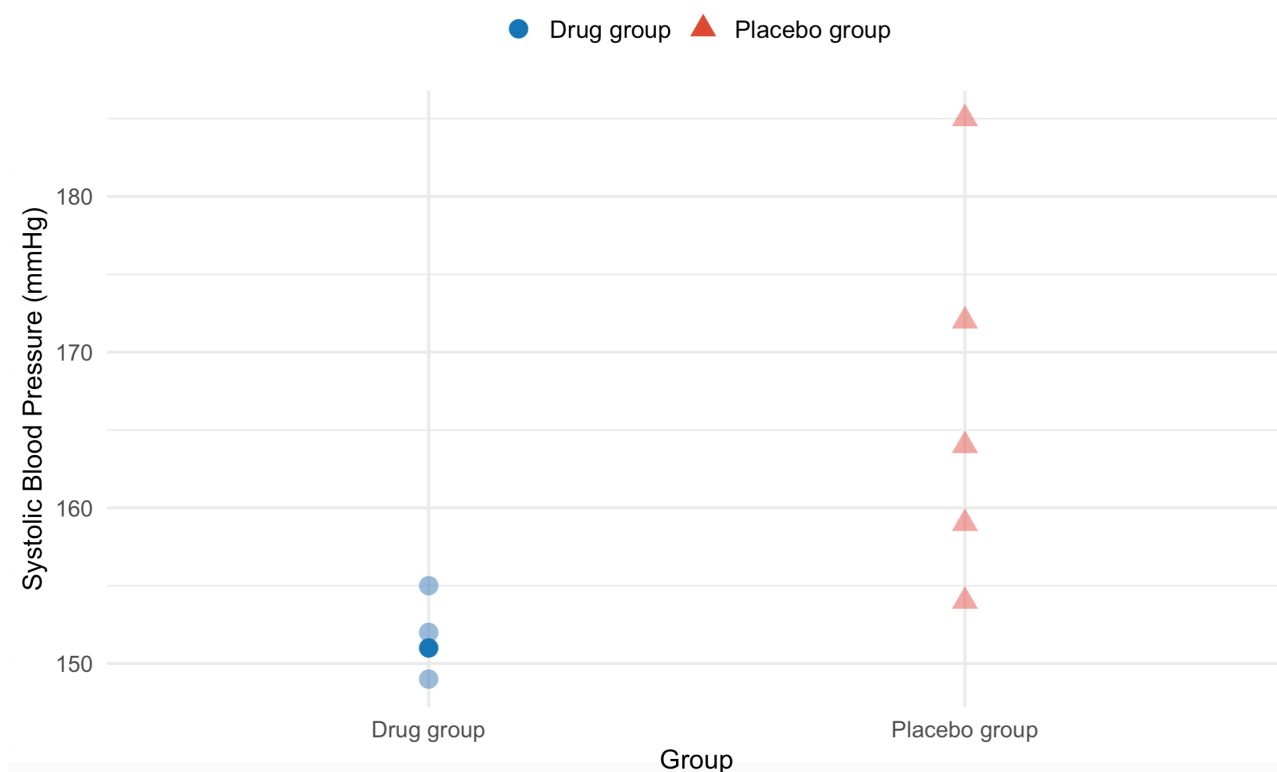


図 2: 両群 2 か月後の収縮期血圧の散布図

R 言語の ggplot を用いて、横軸を群、縦軸を 2 か月後の収縮期血圧とした散布図を図 2 のように作成した。図 2 に示すように、青色の点と赤色の三角形はそれぞれ薬群 (Drug group) とプラセボ群 (Placebo group) の 2 か月後の収縮期血圧を表している。ここで、薬群のデータにおいては、2 名の被験者が同一 151mmHg の収縮期血圧が測定したため、重なることがためにより濃い青色の点で表示している。グラフから分かるように、薬群の収縮期血圧はプラセボ群と比較して全体的に低く、分布のばらつきも小さくなる傾向が見られる。この結果から、2 つの群の間に治療効果の差が存在する可能性が示唆されるが、この差が統計的に有意であるかどうかについては、さらなる統計的な検定による確認する必要があると考える。

課題4の結果

両群の「age」, 「sbp b」, 「dbp b」, 「sbp a」, 「dbp a」, 「sex」の度数を統計し、以下の図3のようなヒストグラムを作成した。

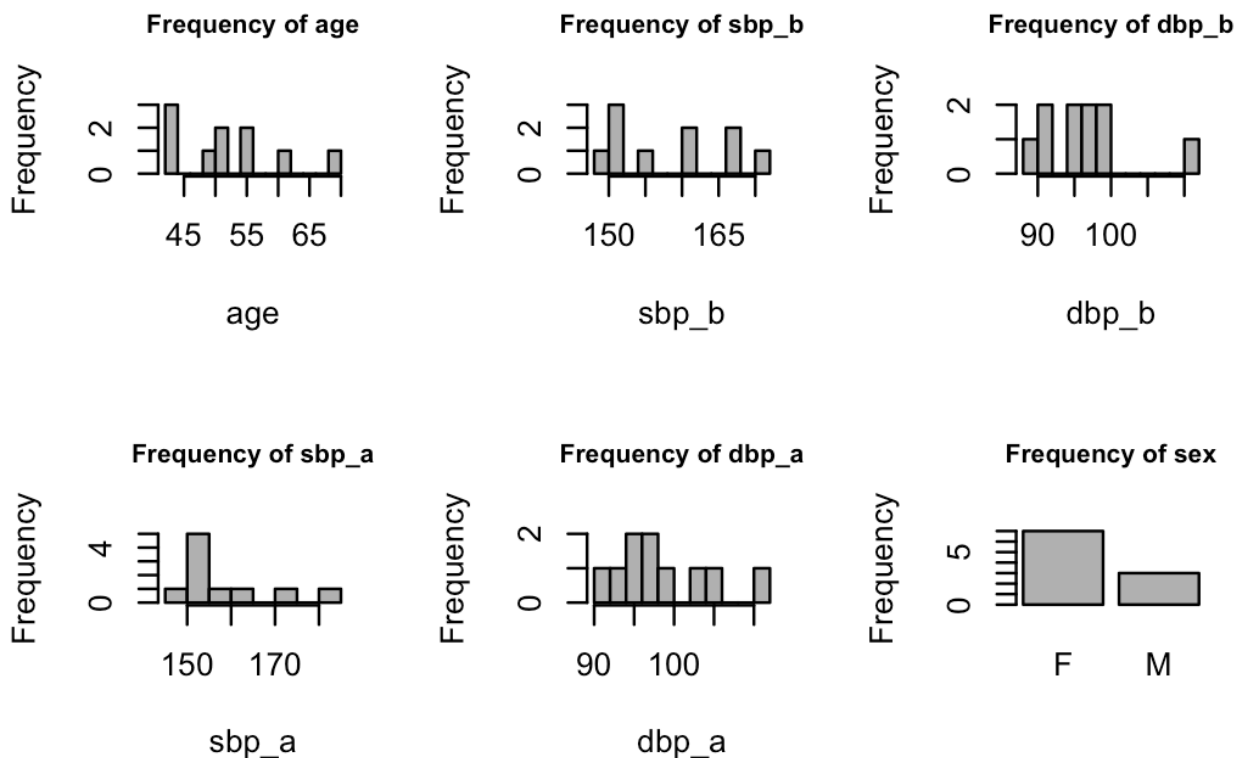


図 3: 群ごとのヒストグラム

課題1で算出したデータと度数の分布に基づいて、まず、年齢に関しては、開始時点薬群とプラセボ群の平均年齢は約13歳の差が見られ、収縮期血圧では、プラセボ群の方が約14 mmHg高かった。拡張期血圧では薬群96.0 mmHg、プラセボ群97.6 mmHgとほぼ同じ程度であり、性別の構成では課題1で述べたように、実験参加者の数が少ないため、薬群の女性がやや多い(より一人だけ)傾向が確認された。

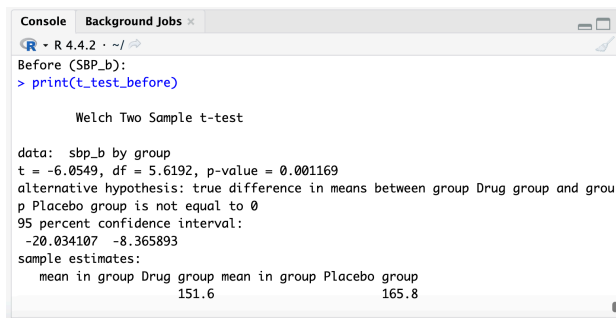
これらの結果から、開始時点における年齢および収縮期血圧(SBP(b))に群間差が存在し、2群のバランスは十分でないと考えている。このような不均衡は、治療効果の推定に交絡をもたらす可能性がある。そのため、実際の臨床データを処理する時、具体的には、年齢や基準時血圧、性別などの共変量を考慮する方が良いと考えている。例えば、回帰分析や相関分析により、他の変数の影響を取り除いたうえで群の効果を評価したり、データの量が多い場合マッチング法を導入することにより、擬似的にランダム化に近い条件を再現し、群間比較の精度を高めることが可能だと考えられる。

課題 5 `t.test` 関数を使用して、開始時点と 2 か月後の時点ごとに、2 群間の収縮期血圧の母平均が同じかどうか、 t 検定で調べてください。さらに、母平均の差の 95% 信頼区間の値を確認してください。

課題 6 「sbp_a」と「sbp_b」の差を表す変数「diff」を作成して、2 群間の収縮期血圧の変化量の母平均が同じかどうか、 t 検定で調べてください。さらに、母平均の差の 95% 信頼区間の値を確認してください。

課題 5 の結果

`t.test` 関数を用いて、開始時点 (SBP(b)) および 2 か月後 (SBP(a)) における薬群とプラセボ群の収縮期血圧の母平均を比較し、結果は図 4 のように示される。

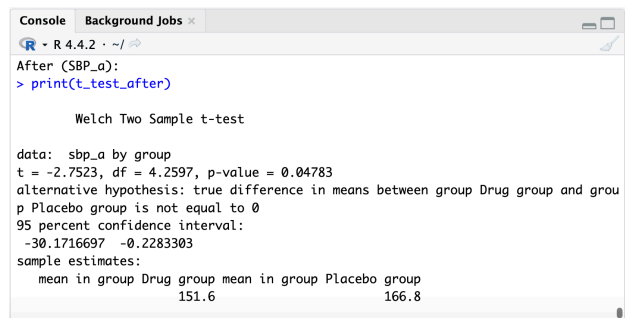


```
Console Background Jobs
R 4.4.2 ~
Before (SBP_b):
> print(t_test_before)

Welch Two Sample t-test

data: sbp_b by group
t = -6.0549, df = 5.6192, p-value = 0.001169
alternative hypothesis: true difference in means between group Drug group and group
p Placebo group is not equal to 0
95 percent confidence interval:
-20.034107 -8.365893
sample estimates:
mean in group Drug group mean in group Placebo group
151.6 165.8
```

(a) 開始時点の収縮期血圧の比較の t 検定結果



```
Console Background Jobs
R 4.4.2 ~
After (SBP_a):
> print(t_test_after)

Welch Two Sample t-test

data: sbp_a by group
t = -2.7523, df = 4.2597, p-value = 0.04783
alternative hypothesis: true difference in means between group Drug group and group
p Placebo group is not equal to 0
95 percent confidence interval:
-30.171697 -0.228303
sample estimates:
mean in group Drug group mean in group Placebo group
151.6 166.8
```

(b) 2 か月後の収縮期血圧の比較の t 検定結果

図 4: 課題 5 の結果出力

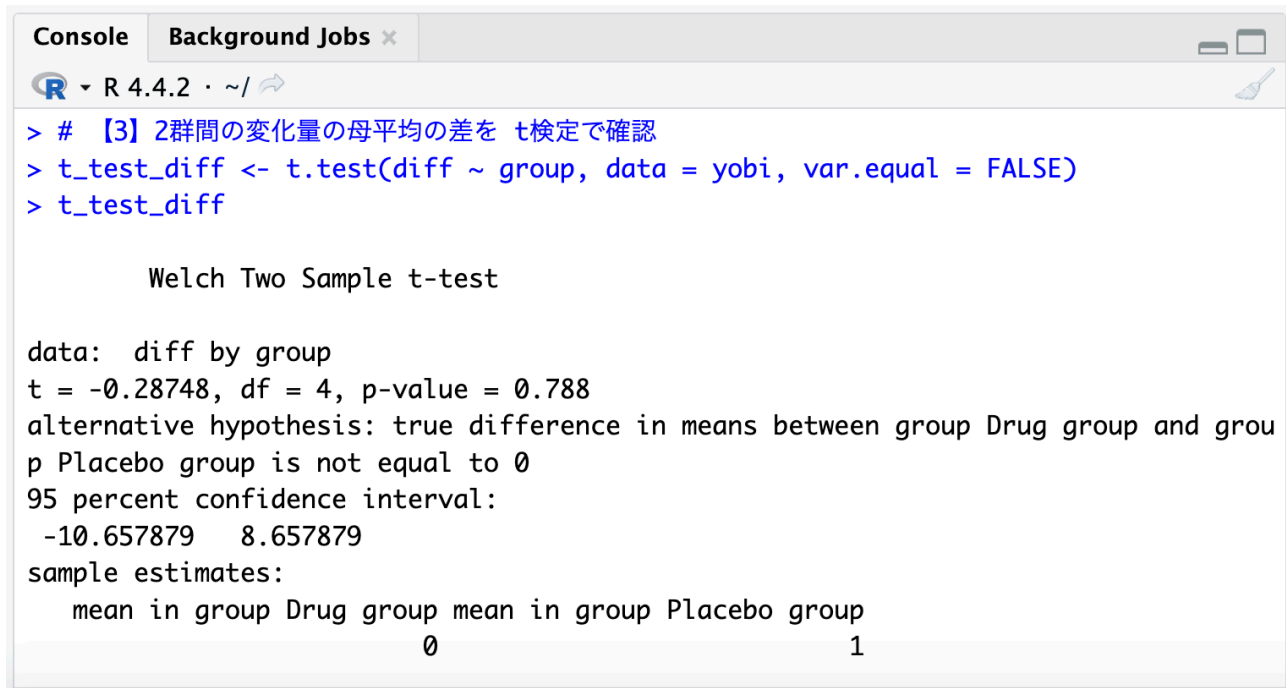
図 4 の結果より、開始時点では $t = -6.05$, $p = 0.0012$ であり、母平均の差の 95% 信頼区間は $[-20.03, -8.37]$ であった。 $p < 0.05$ 且つ信頼区間に 0 が含まれていないため、開始時点の両群の母平均は統計的に有意に異なると示唆される。

また、2 か月後の測定では $t = -2.75$, $p = 0.0478$, 95% 信頼区間は $[-30.17, -0.23]$ であり、こちらも $p < 0.05$ 且つ信頼区間は 0 を含まないため、有意差が認められ、つまり薬群の平均収縮期血圧 (151.6 mmHg) はプラセボ群 (166.8 mmHg) より低く、治療効果の存在が示唆される。

したがって、両時点とも群間に統計的有意差が存在し、特に 2 か月後も薬群で収縮期血圧が有意に低下していることが確認された。しかし、2 群間でベースラインが十分に均衡していなかったことが明らかとなった。この不均衡は、治療効果の推定に交絡をもたらす可能性があるため、実際の実験と解析では基準時点の血圧を共変量として調整することが望ましい。

課題 6 の結果

課題5のように、`t.test` 関数を用いて、「sbp a」と「sbp b」の差として変数「diff」とし、2群の収縮期血圧の母平均を比較し、結果は図5のように示される。



```
Console Background Jobs x
R 4.4.2 ~/
> # 【3】 2群間の変化量の母平均の差を t検定で確認
> t_test_diff <- t.test(diff ~ group, data = yobi, var.equal = FALSE)
> t_test_diff

Welch Two Sample t-test

data: diff by group
t = -0.28748, df = 4, p-value = 0.788
alternative hypothesis: true difference in means between group Drug group and group Placebo group is not equal to 0
95 percent confidence interval:
 -10.657879 8.657879
sample estimates:
 mean in group Drug group mean in group Placebo group
                        0                        1
```

図 5: 両群の変化量の t 検定結果

図5の結果より、 $t = -0.29$, $p = 0.788$ であり、95%信頼区間は $[-10.66, 8.66]$ であった。信頼区間に0が含まれており、 $p > 0.05$ であることから、両群の血圧変化量の母平均に統計的に有意な差は認められなかった。すなわち、薬群とプラセボ群の平均変化量は偶然による変動の範囲内と考えられる。したがって、本データに基づく限り、薬剤による治療効果を統計的に確認することはできなかった。本データでは各群のサンプルサイズが $n = 5$ と非常に小さいため、統計的検出力 (power) が低く、実際に平均値の差が存在しても有意差として検出されにくい可能性がある。したがって、本結果から治療効果が存在しないと断定することはできず、今後はより多くの被験者を対象とした再検証が必要である。

課題 7 主要変数として「2か月後の収縮期血圧の値」と「収縮期血圧の変化量」のどちらを使用するかを決めてください。レポートには、どちらをどのように選んだかを記載してください。予備試験で得られた値を参考にして、被験薬群とプラセボ群の母平均の差 (Δ) と主要変数のばらつきの大きさ (σ) を決めてください。このとき、予備試験の参加者数はわずか 10 人ですから、そこで得られた平均値の差や標準偏差の精度は高いことに注意して。

Δ , σ , 検出力の値をいろいろ変えると、どの程度標本サイズが変わるかを検討しながら、試験に必要な各群の標本サイズ $n/2$ を決めてください。第 1 種の過誤確率は両側 5% とします。必要以上に標本サイズを大きく見積もるべきではありません。各群の標本サイズ ($n/2$) は最大 45 人 (2 群で 90 人) とします。

試験に必要な標本サイズの計算には、例えば `power.t.test` 関数を使用します。レポートには、標本サイズ設計に使用した (1) 期待される大きさ, (2) 標準偏差, (3) 目標とする検出力を記載してください。

課題 8 課題 1 と同様に、連続量のデータの基本統計量 (データ数, 平均, 標準偏差, 中央値, 最小値, 最大値) を群ごとに計算して、結果をみやすい表にして整理してください。2 値変数のデータ「sex」の各水準の度数と割合を群ごとに計算してください。

課題 9 課題 8 の結果から、開始時点の収縮期血圧, 年齢, 性別が, 2 つの群間でバランスがとれているかを確認してください。

課題 10 2 群間の主要変数の母平均が同じかどうか, t 検定で調べてください。さらに, 母平均の差の 95% 信頼区間の値を確認してください。最後に, 得られた母平均の差が臨床的に意味があるかどうかを検討してください。

課題 11 本試験の目的が達成できたかどうかを班で話し合ってください。

課題7の結果

グループメンバーたちが相談した結果、薬の効果を正確に評価するためには、単に投与後の血圧値を見るだけでなく、投与前から投与後までの変化を比較することが重要であると結論つけた。そのため、「収縮期血圧の変化量」を主要変数として採用し、投与前後の変化量を指標とすることで、個人差や初期値のばらつきの影響を抑え、より直接的に薬剤の効果を反映できると考えている。また、主要なパラメータの設定値を以下の表に示す。

項目	設定値
母平均の差 (Δ)	8
標準偏差 (σ)	13
第1種の過誤確率 (有意水準)	両側 5%
検出力 (power)	0.8
合計サンプルサイズ (n)	86 人

課題8の結果

付録のコード（課題8の部分）を実行し、結果をまとめて表4と5のように示す。ここで、

表 4: 群別の基本統計量

グループ	データ数	平均	標準偏差	中央値	最小値	最大値
Age						
Drug group (N)	43	54.49	7.96	54	40	68
Placebo group (P)	43	53.98	8.71	54	40	69
SBP(b)						
Drug group (N)	43	160.67	7.87	161	141	176
Placebo group (P)	43	159.35	9.14	158	139	181
SBP(a)						
Drug group (N)	43	152.84	10.30	153	128	173
Placebo group (P)	43	161.09	9.01	163	143	176
Diff (= SBP(a) - SBP(b))						
Drug group (N)	43	-7.84	13.17	-9	-38	23
Placebo group (P)	43	+1.74	12.53	0	-25	27

SBPの前後変化量(Diff)に注目すると、Drug群では平均変化量が-7.84 mmHg、Placebo群では+1.74 mmHgであり、被験薬群の方が明らかに血圧低下していることを示した。これは被験薬の降圧効果が示唆されていると考えられるが、この差が統計的に有意であるかどうかについては、さらなる統計的な検定による確認する必要があると考える。

表 5: 群別の性別分布

Group	Sex	Count	Percentage (%)
Drug group (N)	F	23	53.5
	M	20	46.5
Placebo group (P)	F	19	44.2
	M	24	55.8

課題 9 の結果

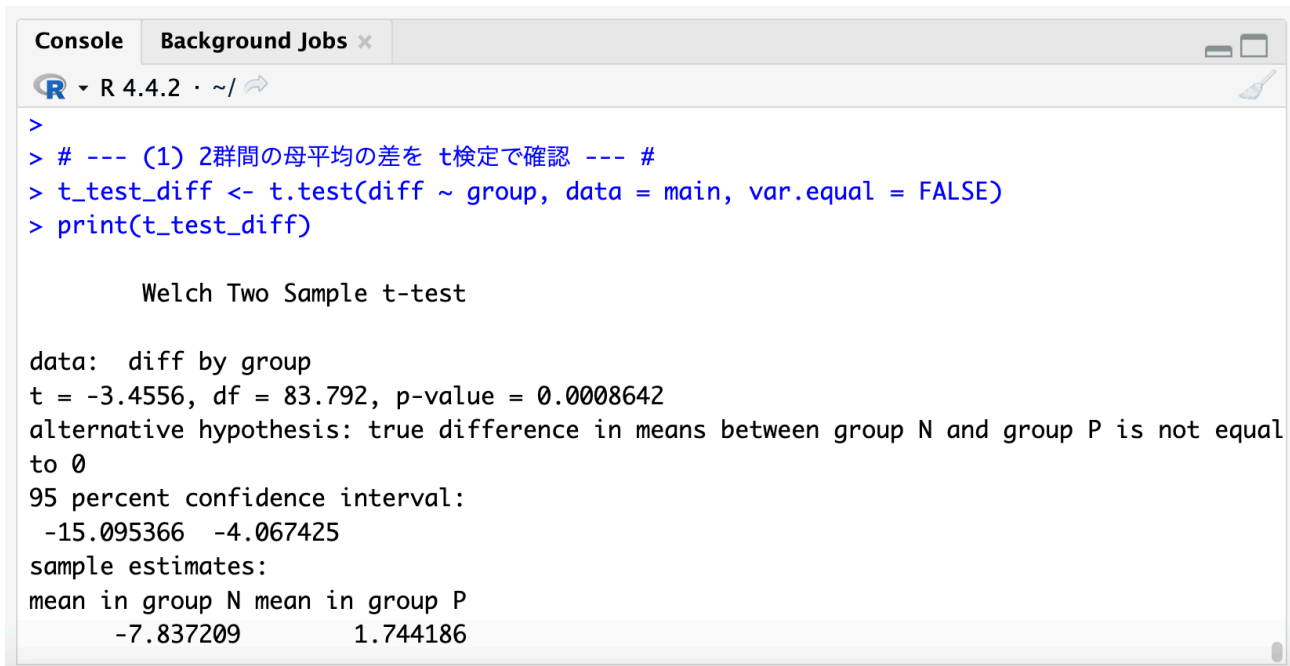
課題 8 の結果より、両群の年齢（Drug 群 54.5 歳、Placebo 群 54.0 歳）には差があまりなく、開始時の収縮期血圧（Drug 群 160.7 mmHg、Placebo 群 159.3 mmHg）は同程度であったことがわかる。一方、表 5 の結果より、性別構成も、Drug 群では女性 53.5%、男性 46.5%、Placebo 群では女性 44.2%、男性 55.8%と、全体的に均衡していることが確認された。

したがって、課題 8 においては、開始時点の年齢、血圧、性別のの偏りがほぼなく、2 群間のバランスは保たれていると考える。

また、課題 4 の結果と比較すると、被験者数の増加により、実験前後両群の年齢、収縮期血圧、性別構成のどちらもバランスがよくなり、不均衡の現象が軽減してきた。

課題 10 の結果

付録のコード（課題 10 の部分）を実行し、結果を図 6 のように示す。



```

>
> # --- (1) 2群間の母平均の差を t検定で確認 --- #
> t_test_diff <- t.test(diff ~ group, data = main, var.equal = FALSE)
> print(t_test_diff)

Welch Two Sample t-test

data:  diff by group
t = -3.4556, df = 83.792, p-value = 0.0008642
alternative hypothesis: true difference in means between group N and group P is not equal to 0
95 percent confidence interval:
 -15.095366  -4.067425
sample estimates:
mean in group N mean in group P
    -7.837209      1.744186

```

図 6: 課題 10 の両群の変化量の t 検定結果

図 6 の結果より、受験者 ($n = 86$) のデータを用いて主要変数である収縮期血圧の変化量 ($\text{Diff} = \text{SBP(a)} - \text{SBP(b)}$) について、2 群間の母平均の差を Welch の t 検定を行った結果は、 $t = -3.46$, 自由度 $df = 83.79$, $p = 0.00086$ であった。ここで、95%信頼区間は $[-15.10, -4.07]$ であったため、信頼区間に 0 が含まれず、 $p < 0.05$ であることから、2 群間の母平均には統計的に有意な差が認められたと考えている。

また、薬群 (Drug group) の平均変化量は -7.84 mmHg 、プラセボ群 (Placebo group) は $+1.74 \text{ mmHg}$ であり、薬群の方が平均して約 9.6 mmHg の血圧低下を示し、課題 8 で設定した母平均の差 $\Delta = 8$, 標準偏差 $\sigma = 13$, 有意水準 $\alpha = 0.05$, 検出力 0.8 に基づいて想定した効果量とほぼ一致している。

平均差が約 9 mmHg と臨床的にも意味のある程度であることから、被験薬はプラセボと比較して実際に有意な降圧効果を有すると考えられる。

したがって、群のサンプルサイズ $n/2 = 43$ という設計条件のもとで、理論的に想定した降圧効果が統計的にも実証されたといえると考ええる。

課題 11 の結果

1 班の班内での討論の結果により、本試験の目的は確かに達成されたと考える。本研究では、薬群とプラセボ群の被験者をそれぞれ $n = 43$ 名ずつ設定し、主要変数として収縮期血圧の変化量 (投与前後の差) を採用した。この設計は、個人の初期血圧や年齢などによるばらつきの影響を抑えつつ、薬剤の効果をより直接的に評価することを目的とする。

その結果は、薬群では平均して約 -7.8 mmHg の血圧低下が観測され、プラセボ群の $+1.7 \text{ mmHg}$ と比較して有意な差 ($t = -3.46$, $p = 0.00086$) が認められた。事前に想定した母平均差 $\Delta = 8 \text{ mmHg}$, 標準偏差 $\sigma = 13$, 検出力 0.8 の条件と整合する結果であり、統計的と臨床的に効果が確認されたといえる。

また、課題 4 の小規模データでは群間に偏りが存在したが、本試験を通じて、事前の分布バランス (年齢、性別、基準時点の血圧など) の均衡を取ることの重要性を認識した。サンプルサイズを十分に確保した本試験では、年齢・性別構成・基準血圧のいずれもほぼ均等となり、このことが統計的検出力の向上に寄与されたと考える。

さらに、標本サイズ設計の重要性についても確認できた。事前に適切な Δ , σ , および検出力を設定してサンプル数を計算したことにより、データ収集の負担をかけすぎずに、良い結果を得ることができた。もし標本数が少なければ、第二類過誤のリスクが高まり、本試験のような明確な結論は得られなかったと考えられる。

以上より、本試験の設計・解析は妥当であり、薬剤の降圧効果が統計的に確認されたことから、研究目的は達成されたと判断する。また、今後より大規模な試験を実施する際にも、事前分布の均衡化と標本サイズ設計を重視することが、信頼性の高い臨床評価につながると考えられる。

付録

```
1 # =====
2 #   Data Loading and Preparation
3 # =====
4
5 getwd()
6
7 base0 <- read.csv("/Users/fengmc/Library/CloudStorage/OneDrive東京理科大学-/大
  学/課題/3/Experiment_Souzu/6/base0.csv")
8 after0 <- read.csv("/Users/fengmc/Library/CloudStorage/OneDrive東京理科大学-/大
  学/課題/3/Experiment_Souzu/6/after0.csv")
9
10 # Merge the two datasets by ID
11 yobi <- merge(base0, after0, by = "id")
12
13 # Convert group variable to factor with English labels
14 yobi$group <- factor(yobi$group,
15                       levels = c("N", "P"),
16                       labels = c("Drug group", "Placebo group"))
17
18 # Check the data
19 head(yobi)
20
21 # =====
22 #   Task 1: Descriptive Statistics
23 # =====
24 library(dplyr)
25
26 # ---- (1) Summary statistics for continuous variables ----
27 summary_stats <- yobi %>%
28   group_by(group) %>%
29   summarise(
30     n = n(),
31     age_mean = mean(age), age_sd = sd(age),
32     age_median = median(age), age_min = min(age), age_max = max(age),
33     sbp_b_mean = mean(sbp_b), sbp_b_sd = sd(sbp_b),
34     sbp_b_median = median(sbp_b), sbp_b_min = min(sbp_b), sbp_b_max = max(
35       sbp_b),
36     dbp_b_mean = mean(dbp_b), dbp_b_sd = sd(dbp_b),
37     dbp_b_median = median(dbp_b), dbp_b_min = min(dbp_b), dbp_b_max = max(
38       dbp_b),
39     sbp_a_mean = mean(sbp_a), sbp_a_sd = sd(sbp_a),
40     sbp_a_median = median(sbp_a), sbp_a_min = min(sbp_a), sbp_a_max = max(
41       sbp_a),
42     dbp_a_mean = mean(dbp_a), dbp_a_sd = sd(dbp_a),
43     dbp_a_median = median(dbp_a), dbp_a_min = min(dbp_a), dbp_a_max = max(
44       dbp_a)
45   )
46 summary_stats
47
48 # → Displays mean and SD for each group
49
50 # ---- (2) Histograms (black^^e2^^80^^93gray style) ----
51 num_vars <- c("age", "sbp_b", "dbp_b", "sbp_a", "dbp_a")
52
53 par(mfrow = c(2, 3), bg = "white") # 2x3 layout, white background
54 for (var in num_vars) {
```

```

52 hist(yobi[[var]],
53       main = paste("Frequency of", var),
54       xlab = var,
55       col = "gray70",      # gray bars
56       border = "black",   # black borders
57       breaks = 10,
58       cex.main = 0.8)
59 }
60
61 barplot(table(yobi$sex),
62         main = "Frequency of sex",
63         col = "gray70",
64         border = "black",
65         ylab = "Frequency",
66         cex.main = 0.8)
67
68 # ---- (3) Sex distribution ----
69 sex_summary <- yobi %>%
70   group_by(group, sex) %>%
71   summarise(count = n()) %>%
72   mutate(percentage = count / sum(count) * 100)
73
74 sex_summary
75
76
77 # =====
78 # Task 2: Correlation and Scatter Plots
79 # =====
80
81 # ---- (1) Correlation matrix ----
82 cor_matrix <- yobi %>%
83   select(age, sbp_b, dbp_b, sbp_a, dbp_a) %>%
84   cor(use = "complete.obs")
85 cor_matrix
86
87 # ---- (2) Pairwise scatter plots ----
88 pairs(yobi[, c("age", "sbp_b", "dbp_b", "sbp_a", "dbp_a")],
89       main = "Pairwise Scatter Plots",
90       pch = 21, col = "black", bg = "gray80")
91
92
93 # =====
94 # Task 3: Group Comparison (SBP after)
95 # =====
96
97 library(ggplot2)
98 library(dplyr)
99
100 yobi_plot <- yobi %>%
101   group_by(group, sbp_a) %>%
102   summarise(count = n(), .groups = "drop")
103
104 # プロット
105 ggplot(yobi_plot, aes(x = group, y = sbp_a,
106                       color = group, shape = group)) +
107   geom_point(aes(alpha = count), size = 4, stroke = 1.2) +
108   scale_alpha_continuous(range = c(0.5, 1), guide = "none") +
109   scale_color_manual(values = c("#1F77B4", "#E24A33")) + # ^^e8^^93^^9d & 橙

```

```

111 scale_shape_manual(values = c(16, 17)) + # 三角
112 labs(
113   title = "Systolic Blood Pressure by Group (After 2 months)",
114   x = "Group",
115   y = "Systolic Blood Pressure (mmHg)"
116 ) +
117 theme_minimal(base_size = 16) +
118 theme(
119   legend.position = "top",
120   legend.title = element_blank(),
121   plot.title = element_text(size = 18, face = "bold", hjust = 0.5),
122   axis.title = element_text(size = 14),
123   axis.text = element_text(size = 12),
124   panel.grid.major.y = element_line(color = "grey80", size = 0.4)
125 )
126
127 # =====
128 # Task 4: Group Balance Check
129 # =====
130 #ok
131
132
133 # =====
134 # Task 5: Two-sample t-tests
135 # =====
136
137 # 【 開始時点 (1SBP_) の群間比較b
138 t_test_before <- t.test(sbp_b ~ group, data = yobi, var.equal = FALSE)
139 t_test_before
140
141 # 【 か月後 (22SBP_) の群間比較a
142 t_test_after <- t.test(sbp_a ~ group, data = yobi, var.equal = FALSE)
143 t_test_after
144
145 # 【 3result
146 cat("\n--- t-test Results ----\n")
147 cat("Before (SBP_b):\n")
148 print(t_test_before)
149
150 cat("\nAfter (SBP_a):\n")
151 print(t_test_after)
152
153 # =====
154 # Task 6
155 # =====
156
157 # 【 変化量1
158 yobi$diff <- yobi$sbp_a - yobi$sbp_b # か月後2 - 開始時点
159
160 # 【 2
161 library(dplyr)
162 diff_summary <- yobi %>%
163   group_by(group) %>%
164   summarise(
165     n = n(),
166     mean_diff = mean(diff, na.rm = TRUE),
167     sd_diff = sd(diff, na.rm = TRUE),
168     median_diff = median(diff, na.rm = TRUE),

```

```

170     min_diff = min(diff, na.rm = TRUE),
171     max_diff = max(diff, na.rm = TRUE)
172 )
173 diff_summary
174
175 # 【】 3t
176 t_test_diff <- t.test(diff ~ group, data = yobi, var.equal = FALSE)
177 t_test_diff
178
179
180
181
182 # --- データの読み込み --- #
183 base1 <- read.csv("/Users/fengmc/Library/CloudStorage/OneDrive 東京理科大学-/大
184 学/課題/3/Experiment_Souzu/6/base1.csv")
185 after1 <- read.csv("/Users/fengmc/Library/CloudStorage/OneDrive 東京理科大学-/大
186 学/課題/3/Experiment_Souzu/6/after1.csv")
187
188 # --- ライブラリの読み込み --- #
189 library(dplyr)
190
191 # --- データのソート (と順) groupid --- #
192 base1s <- arrange(base1, group, id)
193 after1s <- arrange(after1, group, id)
194
195 # --- データのマージ (対応付け) --- #
196 main <- left_join(base1s, after1s, by = c("group", "id"))
197
198 # --- 収縮期血圧の変化量 (差分) を計算 --- #
199 main$diff <- main$sbp_a - main$sbp_b
200
201 #####
202 # 課題8
203 #####
204
205 # --- (1) 連続量データの基本統計量 --- #
206 summary_stats <- main %>%
207   group_by(group) %>%
208   summarise(
209     n = n(),
210     age_mean = mean(age, na.rm = TRUE),
211     age_sd = sd(age, na.rm = TRUE),
212     age_median = median(age, na.rm = TRUE),
213     age_min = min(age, na.rm = TRUE),
214     age_max = max(age, na.rm = TRUE),
215
216     sbp_b_mean = mean(sbp_b, na.rm = TRUE),
217     sbp_b_sd = sd(sbp_b, na.rm = TRUE),
218     sbp_b_median = median(sbp_b, na.rm = TRUE),
219     sbp_b_min = min(sbp_b, na.rm = TRUE),
220     sbp_b_max = max(sbp_b, na.rm = TRUE),
221
222     sbp_a_mean = mean(sbp_a, na.rm = TRUE),
223     sbp_a_sd = sd(sbp_a, na.rm = TRUE),
224     sbp_a_median = median(sbp_a, na.rm = TRUE),
225     sbp_a_min = min(sbp_a, na.rm = TRUE),
226     sbp_a_max = max(sbp_a, na.rm = TRUE),

```



```

227
228     diff_mean = mean(diff, na.rm = TRUE),
229     diff_sd = sd(diff, na.rm = TRUE),
230     diff_median = median(diff, na.rm = TRUE),
231     diff_min = min(diff, na.rm = TRUE),
232     diff_max = max(diff, na.rm = TRUE)
233 )
234
235 print(summary_stats)
236
237
238 # --- (2) 値変数 (2) の度数と割合 sex --- #
239 sex_summary <- main %>%
240   group_by(group, sex) %>%
241   summarise(count = n()) %>%
242   mutate(percentage = round(count / sum(count) * 100, 1))
243
244 print(sex_summary)
245
246 #####
247 # 課題9
248 #####
249 #ok
250
251 #####
252 # 課題：主要変数 (10Diff = SBP(a) - SBP(b)) の母平均の比較
253 #####
254
255 # --- (1) 群間の母平均の差を2 検定で確認t --- #
256 t_test_diff <- t.test(diff ~ group, data = main, var.equal = FALSE)
257 print(t_test_diff)

```

Listing 1: 演習用コード