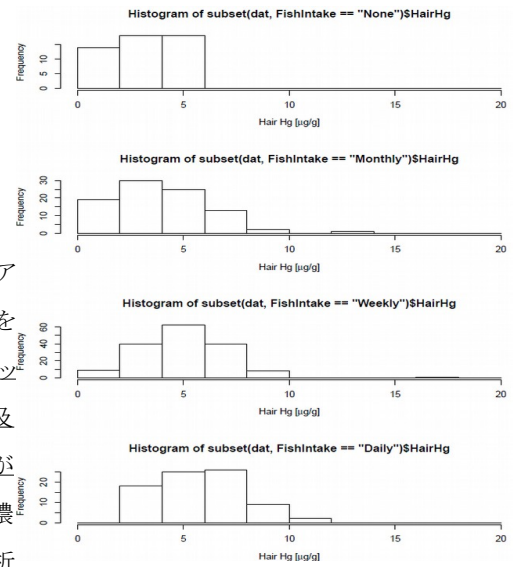


(解答例)

1. 以下の各文の下線部について解釈や解析方法について不適切な点があれば指摘し、改善案を示せ。適切ならば適切と記せ。なお、すべて架空の事例であり、仮に実際に似た状況があったとしても偶然の一致である。

(1) 途上国の海沿いに位置し人口約 3 万 8 千人の R 郡では、数十年前から住民登録が整備されている。Google Earth で見ると同じような規模の村が 50 ほど散在している。近年、魚介類や鯨類を多食する住民で神経症状を示す者が現れ、水銀中毒の可能性が疑われたため、実態調査として、(A)住民登録からランダムに選んだ 1% である 380 人の住民を対象に、魚介類及び鯨類の摂取頻度について、[1]ほとんど食べない、[2]月 1 回以上週 1 回未満、[3]週 1 回以上毎日未満、[4]ほぼ毎日、から選んでもらうと同時に、毛髪水銀濃度を調べたところ、下表の結果が得られた (<http://minato.sip21c.org/fish-Hg-2017.txt> というタブ区切りテキスト形式でアップロード済。変数は PID (個人 ID 番号), HairHg (毛髪水銀濃度), HighHairHg (5 μ g/g 以上なら 1, 未満なら 0), FishIL (上記の魚介類及び鯨類摂取頻度カテゴリ番号) の 4 つ)。

Eating fish or whale frequency	N	Median (Hg 濃度)	Mean \pm SD (Hg 濃度)	毛髪 Hg 濃度 5 μ g/g 以上
1. None	50	3.31	3.14 \pm 1.74	8
2. Monthly	90	3.84	3.92 \pm 2.21	29
3. Weekly	160	4.97	5.04 \pm 2.15	80
4. Daily	80	5.82	5.73 \pm 1.94	52



このデータから魚介類及び鯨類の摂取頻度と水銀曝露の関係を分析するには、2 つのアプローチがある。1 つは、成人耐容摂取量の目安である 5 μ g/g 以上の毛髪中水銀濃度を示すかどうかを魚介類及び鯨類の摂取頻度と独立かを検定する方法である。(B)フィッシャーの直接確率検定をすると、 $p < 0.001$ となって帰無仮説は棄却されるので、魚介類及び鯨類の摂取頻度と毛髪中水銀が 5 μ g/g 以上の高値を示すかどうかには有意な関連があるといえる。もう 1 つのアプローチは、魚介類及び鯨類の摂取頻度の違いが毛髪水銀濃度に影響しているかどうかを調べる方法である。(C) Welch の方法による一元配置分散分析の結果は、F 値 25.591、第一自由度 3、第二自由度 163.13 となるので、やはり $p < 0.001$ となって帰無仮説は棄却される。さらに魚介類及び鯨類の摂取頻度が異なるどの群とどの群の間で毛髪水銀濃度に差があるのか総当たりで調べるには、Welch の t 検定を繰り返せば良い。

(A) 間違いではないが 38,000 人からランダムサンプルした 380 人を一人ずつ訪ねて調査を行うのは時間も手間もかかり過ぎるので、50 村落から 3~5 の村落をランダムに選んで全数調査するクラスターサンプリングでも代表性はあるので、その方が効率が良い。

(B) 正しい。

(C) 4 群の中で 2 組ずつの比較を総当たりで繰り返すと検定の多重性の問題が起こるので不適切。Holm の方法か FDR 法で検定の多重性を調整するか、あるいは Tukey の方法で同時信頼区間を求める必要がある。

(2) 日本において、血圧低下作用が期待される食品 A の効果を調べる目的で、5 人の高血圧患者に A を半年間毎日食べて貰うだけの介入実験をしたところ、開始前と半年経過後の 2 時点間で収縮期血圧(mmHg)が 160→145、150→125、170→155、155→135、145→130 と変化した。対応のある t 検定をすると $p = 0.0008$ となるので、A には統計的に有意な血圧低下効果があると結論できる。

日本で半年経過すると季節変化等時間経過に伴う他の因子の変化の影響を受けて血圧変化する可能性があるため、患者をランダムに A の摂取を義務づけけない群にも割り付け、2 群間で血圧変化を比較するデザインにしなくては血圧変化が A の摂取によるとは言えない。また途中経過も調べるべき。拡張期血圧も調べる方がよい。(注)5%水準で有意差が出ているので検出力には問題がなく、サンプルサイズが小さいこと自体は OK

(3) マーカーの測定値がある閾値を超えると疾病 X を疑うべきとされる標準的な定量的検査方法 A に対して、より安価または迅速等の理由で、同じマーカーの新しい検査方法 B を開発したとする。B が妥当であることを示すには、十分な数の患者と健常者に対して A と B での測定値を対応のある t 検定で比較して有意差がなければ良い。

差の平均がゼロであるだけでは不十分で、相関も高くなければならず、値の大小に伴う系統的なズレがないことも確認せねばならないので Bland-Altman プロットで一致を確認する必要がある。その上で、A と B で描いた ROC 曲線に統計的な有意差がないことも確認するべき。

(4) 慢性肝炎患者をランダムに 22 人ずつの 2 群に分け、片方はプレドニソロンを投与し、残りは経過観察したところ、観察終了時にプレドニソロン群では 11 人が生存しており、経過観察群では 6 人が生存していたが、フィッシャーの直接確率検定の結果、p 値は 0.215 で観察期間終了後の生存確率に有意差はなかった。死亡または観察打ち切りまでの月数のデータ (<http://minato.sip21c.org/hepatitis-2017.txt> にアップロードしており、変数は time(死亡または観察打ち切りまでの月数)、flag(観察終了時死亡が 1、生存が 0)、group(1 がプレドニソロン群、2 が経過観察群)の 3 つである)があるので、死亡例のみについて、プレドニソロン群 11 人と経過観察群 16 人の生存月数の平均値を Welch の t 検定で比較したところ、プレドニソロン群が 80 ヶ月、経過観察群が 31.5 ヶ月で、p 値が 0.02 だったため、有意水準 5% で考えると、プレドニソロンには統計的に有意な延命効果があると結論できた。

死亡例のみにデータを限定すると、相対的に長生きした観察打ち切りデータを除外することになるため、生存時間を過小評価するバイアスが生じる。全データを用い、 Kaplan-Meier 法で生存時間の中央値を推定するとプレドニソロン群が 146 ヶ月、経過観察群が 40.5 ヶ月となり、ログランク検定をすると、 $p=0.0309$ なので、プレドニソロン群は統計的に有意な延命効果があると結論できる(注:このデータでは偶々結論は同じだが、用いる統計手法が不適切)。

2. 集団における疾病量を示す値としての有病割合(prevalence proportion)について説明せよ。この値を得るためにどういうデザインの研究がされるかにも触れること。

有病割合とは、横断研究(断面的研究, cross sectional study)をした際に、対象者全員の中で、その疾病にかかっている人の割合を示す値であり、その集団への疾病負荷(disease burden)の大きさを意味する。

3. 10 人の健常被験者について糖負荷試験を行い、負荷前、負荷直後、30 分後、1 時間後、2 時間後、3 時間後の 6 時点で血中無機リン酸濃度を測定した。測定値が経時的に変化するかどうか検定する統計手法としては何が使えるか説明せよ。

フリードマンの検定も使えるが、極端な外れ値がなければ反復測定分散分析を行うのが最適である。

4. 紅茶を飲むと短期記憶が向上するかどうかを調べるため、10 人の被験者に対して、紅茶を飲む前後でフラッシュ記憶をしてもらって得点を比較するという実験研究を行った。結果が下表の通りだったとき、紅茶を飲んで短期記憶は向上したと言えるか? 有意水準 5% で検定せよ (EZR や関数電卓を使ってもよいが、自由度 9 の t 分布の 97.5% 点が 2.262 であることと $\sqrt{2}=1.414$, $\sqrt{3}=1.732$, $\sqrt{5}=2.236$ のどれかを使えば四則演算だけで計算できる)

紅茶飲用前の得点	7	8	7	9	3	7	5	8	7	6
紅茶飲用後の得点	9	9	7	10	4	10	6	9	8	10

R で $t.test(c(7, 8, 7, 9, 3, 7, 5, 8, 7, 6), c(9, 9, 7, 10, 4, 10, 6, 9, 8, 10), paired=TRUE)$ を実行すると、 $p=0.003$ となるので、変化しないという帰無仮説は棄却される。紅茶を飲んだ後で、統計的に有意に短期記憶が向上したと言える。

手計算の場合は、引用後の得点から引用前の得点を引いた値が 2, 1, 0, 1, 1, 3, 1, 1, 1, 4 なので平均 1.5、各々から 1.5 を引いて 2 乗すると 0.25, 0.25, 2.25, 0.25, 0.25, 2.25, 0.25, 0.25, 0.25, 6.25 となるので、その和の 12.5 を 9 で割った値が不偏分散。

t 値は $1.5/\sqrt{(12.5/9/10)}=9/\sqrt{5} > 4 > 2.262$ なので有意水準 5% で帰無仮説は棄却される。以下同様。

5. 腹痛に対して従来薬と新薬での頓服治療をランダムに割り付ける RCT で、新薬が従来薬よりも優れた効果をもつかを調べたいとする。従来薬の服用では 75% の人に腹痛軽減効果があるとされている。新薬で効果が見られる人の割合が 10% 以上増えれば臨床的に意味があると考えられるとして、有意水準 5%、検出力 80% で **両側カイ二乗検定** を行いたい場合、必要なサンプルサイズは、従来薬群と新薬群が同数として何人ずつか。

$power.prop.test(p1=0.75, p2=0.85, sig.level=0.05, power=0.8)$

より、各群 250 人

*うりぼーネットでの授業評価をしてください。忌憚のないご意見をお願いします。