

(裏面もあるので注意。解答例は<http://minato.sip21c.org/lastminitest-2014-ans.pdf>に公開予定)

領域 \_\_\_\_\_ 学籍番号 \_\_\_\_\_ 氏名 解答例

1. 以下の各文について解釈や解析方法の誤りを指摘し、どうすれば良いか提案せよ

(1) Q市の住民の5%サンプルである1,000人の住民について、食塩摂取量と収縮期血圧の相関係数を求め、母相関係数が0という帰無仮説の検定をしたところ、 $r=0.12$ ,  $p=0.000154$ だった。 $p \ll 0.05$ なので、Q市では食塩摂取と収縮期血圧の間に明らかに強い相関があるといえる。

有意だが極めて弱い相関 (注: p値が小さいことは母相関係数がゼロの時これより絶対値が大きいrが得られる確率の小ささを示すのみ)

(2) 化学工場の爆発事故が起こってヒトへの毒性が報告されている(けれども自然にも微量は存在している)物質Aが飛散し、一般住民が曝露したとする。曝露した住民と、工場から離れた場所に居住している非曝露群の住民を(性別・年齢を曝露した住民とマッチングして)15人ずつサンプリングし、採血して、血液中のAの濃度を比較したところ、曝露群で平均12.5、標準偏差2.3、非曝露群で平均10.5、標準偏差3.4であり、Welchの方法によるt検定の結果、 $t_0=1.887$ ,  $\phi=24.6$ ,  $p=0.071$ であった。検定結果が有意水準5%で有意でなかったため、爆発事故の影響はないと結論できる。

があるとはいえない (注: 有意でないときは帰無仮説は保留。とくにサンプルサイズが小さいので検出力が小さい可能性が高い)

(3) 目的変数Zを説明変数P, Q, Rで説明する重回帰分析の結果、P, Q, Rの偏回帰係数の中でPの係数の絶対値が最大だったので、この3つの変数の中でZへの影響が最も大きいのはPである。

標準化偏回帰係数 (注: 偏回帰係数そのものでは、変数そのものスケールの影響を受けてしまうため、影響の大きさは評価できない)

(4) 化学療法によって寛解した白血病患者40人から同意を得て、ランダムに2群に分け、20人は新薬Xにより維持化学療法を実施し、他方の20人は従来薬により維持化学療法を実施して、再発までの時間に差がないかどうかを分析したところ、新薬群では6名、従来薬群では3名が研究期間内に再発しなかったため対象から除き、新薬群14名と従来薬群17名の間で再発までの時間をWelchの方法による平均値の差のt検定をしたところ、p値が0.06だったので、有意水準5%で考えると、新薬と従来薬の薬効に統計的に有意な差はないといえる

これらの人を打ち切りデータとして、ログランク検定または一般化ウィルコクソン検定またはコックス回帰など、生存時間解析を行う必要がある (注: 除いてしまうと再発までの時間を過小評価することになる)

2. 集団における疾病量を示す値としての罹患率(incidence rate)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必要かにも触れること。

原則としてコホート研究において得られる。観察人時で罹患患者数を割って得られ、疾病の発生速度(感染症ならば伝播の速さ)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時点での血糖値を測定したとき、糖負荷後の血糖値が経時的に変化するかどうか検定する統計手法について、簡潔に説明せよ。

時点ごとの血糖値を変数として、反復測定分散分析(Repeated Measures ANOVA)で時間の効果の有意性検定を行えばよい (注: フリードマンの検定でも良い。いずれにせよ10人の一人一人の血糖値の変化を線で結ぶグラフ表示は必須)

4. 珈琲を飲むと計算能力が向上するかどうかを調べるため、8人の被験者に対して、珈琲を飲む前後で単純計算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。

(1) 適切な検定方法と帰無仮説について簡潔に説明せよ。

珈琲を飲む前後の得点の差の平均値がゼロという帰無仮説に対して、一変量 t 検定を行う（または、得点に差がないという帰無仮説を検定する対応のある t 検定を行う）

(2) 結果が下表の通りだったとき、珈琲を飲んで計算能力は向上したと言えるか？ 有意水準 5%での検定結果を示せ。なお、1 標本 t 検定の t 値はサンプルの平均値から想定される母平均を引いた値を、サンプルの標準誤差（つまりサンプルの不偏標準偏差をサンプルサイズの平方根で割った値）で割って得られる。また、自由度 9 の t 分布の 97.5%点は 2.262,  $\sqrt{5}=2.236$  とする。（←これでは情報不十分でした）

自由度 7 の t 分布の 97.5%点は 2.364624,  $\sqrt{2}=1.414$ ,  $\sqrt{5}=2.236$ ,  $\sqrt{7}=2.646$ ,  $\sqrt{11}=3.317$  とする。

珈琲飲用前の得点	4	7	8	8	6	4	5	6
珈琲飲用後の得点	5	8	10	9	7	6	4	8
得点の差	1	1	2	1	1	2	-1	2

差の平均は $(1+1+2+1+1+2-1+2)/8 = 1.125$ ,

不偏分散は $(1/64+1/64+49/64+1/64+1/64+49/64+289/64+49/64)/7 = (440/64)/7 = 55/56$

$t_0=1.125/\sqrt{((55/56)/8)}=1.125*8/(\sqrt{5}*\sqrt{11}/\sqrt{7})=9*2.646/(2.236*3.317)=3.211 > 2.36$  から帰無仮説が棄却され、有意に得点が向上したといえる。

5. マレーシアでは、従来、 Dengue 熱の報告は発熱などの臨床症状によっていたが、チクングニアなど、似た症状の他の病気が同時に流行しているため、特異度が低いことが問題視されている。ある村の診療所で、ある日に発熱等で外来した子供 40 人について、臨床症状による診断と同時に指先採血サンプルを用いて ELISA 法で確定診断したところ、ともに Dengue 熱と診断された子供が 12 人、臨床症状では Dengue 熱と診断されたが ELISA では陰性だった子供が 8 人、臨床症状では Dengue 熱とは判定されなかったが ELISA で陽性だった子供が 4 人、ともに陰性だった子供が 16 人であった。また、その日に診療所に来なかった子供は村に 260 人いて、その子供たちは全員健康だったとする。

(1) その村の**子供全員**におけるその日の Dengue 熱の有病割合を計算せよ。（←問題が不十分でした）

ELISA が確定診断なので、Dengue 熱の患者は  $12+4=16$  人。全体で 300 人なので、有病割合は  $16/300=5.3\%$

(2) 臨床症状による Dengue 熱診断の感度と特異度を計算せよ。

感度は患者のうち臨床症状で陽性となる人数なので  $12/16=0.75$ 、特異度は非患者のうち臨床症状で陰性となる人数なので、 $16/24=0.67$

感度 0.75      特異度 0.67

6. 高血圧患者に対して従来薬と新薬での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。先行研究によると、従来薬を 2 ヶ月服用した群での効果は、SBP が平均 8 mmHg（標準偏差 3 mmHg）低下することであった。新薬が従来薬よりも 5 mmHg 以上優れた血圧降下作用を示せ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 90%で片側 t 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。

（計算には正規分布の 2.5%点が  $-1.96$ 、5%点が  $-1.64$ 、10%点が  $-1.28$  であることを用いよ）

$$\frac{2(-1.64 - 1.28)^2 3^2}{5^2} + \frac{(-1.64)^2}{4} = 6.81 \dots \sim 7 \text{ (人)}$$

