

平成 28 年度 実験計画法

第2回(データ解析の基礎)

4月 18日(月) 8:40～10:10

場所:楠元 4号館3階 パソコン室

担当:加藤一夫

< 標本抽出(sampling) >

一般に、集団の性質は、部分(標本)を調べて全体(母集団)を推測する標本調査から行なわれる。

1. 母集団(population)と標本(sample)

(1) 母集団

1)母集団とユニバース

ユニバース:現実に把握できるかどうかを別にした理想の対象集団

母集団:時間と場所を特定したうえで実際にサンプリング可能な集団

2)母集団の転移(母集団の取り違い)

例えば、8020に関連して高齢者の生活実態や意見を調べる調査で、途中で家族の苦勞や意見を聞き始めると、母集団が高齢者なのかその家族なのかわからなくなってしまふ。このように調査の途中で母集団が変わってしまうことを母集団の転移といい、これが起きると標本理論が適用できなくなる。

3)抽出枠(sampling frame)

母集団から標本を抽出する(確率抽出法)もとなるリスト(住民台帳など)。

(2) 標本の大きさ(sample size)の決定

1)必要標本数の推定

あまり大きな誤差のある標本では困るが、費用が不足するほど大きすぎるとは調査が実施できない。一般に、標本の大きさを決める要因は、第1には標本誤差の許容水準であり、次には費用となる。

次の式から信頼度 95 %における標本の大きさを推定することができる。

$$n = \frac{N}{\frac{(n-1)e^2}{1.96^2 P(1-P)} + 1}$$

ただし、Nは母集団の大きさ

nは標本の大きさ

Pは母比率の予想値

eは信頼度 95 %での標本誤差の許容値

【例題1】 10,000人と推定されるある地域の高齢者母集団から8020対象者の比率を求める調査を行うと仮定して、上の式からnを求めてみる。母比率Pはわからないので、これまでの報告から10%と予想し、また許容できる標本誤差eはあまり大きな誤差は困るので2%以内とする。これらの数値を式に代入すると、795人という結果が得られる。つまり、10,000人の中から795人を無作為に抽出すれば、母比率±2%の範囲から95%の確率で標本を得ることができる。(実際には、これに調査の回収率を加味して考える。)

【演習2】 例題1において、標本誤差を半分にするには標本を何倍にしなければならないか。また、許容できる標本誤差が2%の場合、標本サイズが最も大きくなる予想母比率はいくつか。

2. サンプリングの種類

(1)確率抽出法

母集団に含まれるすべての個体が標本に選ばれる確率を等しくするための操作を確

率抽出という。母集団のリストから等確率で無作為に抽出する方法が確率抽出法である。

1)単純無作為抽出法(simple random sampling)

母集団の個体数分のくじを作り、引き当てた番号の個体を標本とする。通常、くじの代わりに乱数表を利用する。

2)系統抽出法または等間隔抽出法(systematic sampling)

例えば、母集団 10,000 人のなかから 1,000 人を選ぶには、10 人に1人を選ばばよい。母集団リストに並んでいる最初の 10 人の中から1人を選ばばよい、あとは 10 人間隔で抽出するような場合を系統抽出法または等間隔抽出法という。

母集団の配列のなかに何らかの周期性があり、その周期が抽出間隔と同じかその整数倍である場合、特殊な個体だけが選ばれる偏った標本となる恐れがある。

3)層化抽出法(stratified random sampling)

例えば、男女と年齢について母集団の構成と同じ割合(比例割当)の標本を抽出するため、母集団リストを男女の層に分け、さらに、男女の各層を年齢別に並べ替えたうえで単純無作為抽出なり系統抽出を行うものを層化抽出法という。単純無作為抽出法の原理を維持しているが、同じ標本数であれば標本誤差は小さくなる。

4)多段抽出法(multi-stage sampling)

調査対象者の居住する地域が全国や愛知県全体といった広域にわたる場合、抽出の枠が大きすぎて直接標本を選ぶことが難しい。このような場合、例えば、第1段階で調査する地点を選び(第1次抽出単位 primary sampling unit)、第2段階では選ばれた地点の中から個人(第2次抽出単位 secondary sampling unit)を選ぶ。抽出段階が多くなるほど標本誤差は大きくなる。

5)層別2段抽出法(stratified two-stage sampling)

4)において、地域を偏りなく選ぶため、最初に地域を人口規模とか産業構造などの指標を用いて性質別に層に分けておき、各層から地域を抽出した後で、第2次抽出を行う場合、層別2段抽出法と呼ぶ。

(2)非確率抽出法の種類

ホームレスの調査で母集団リストを期待しても無理なように、調査対象によっては確率抽出法が適応できないことがある。また、調査の制約から、その地域住民の標本を歯科医院の受診者から選ぶような場合、無作為な抽出法とはいえない。

非確率抽出法には、確率抽出法の代用として標本の代表性確保をねらったクォータ法 quotasampling と、代表性にこだわらない有意抽出法 purposive sampling や便宜的抽出法 available sampling がある。

3. バイアス

実験や調査などの研究から得られる結果を系統的に歪めるもののこと。バイアスを含む実験や調査の結果の科学的価値は低い、あるいはゼロとされる。

(1)選択バイアス

研究対象として観察する集団が適切に選択されなかったときに起きるバイアスで次のようなものがある。

1)入院バイアス

入院患者を対象にして疫学調査を実施する場合、真にその病気の姿を表していないことがある。例えば、心筋梗塞患者の罹患要因を明らかにする目的で入院患者を研究対象とした場合、発症後数時間以内に死亡する重症例は研究対象に含まれないために比較的軽症な患者の特徴を調べることになる。

2)有病－罹患バイアス

症例対照研究で選定する症例群は、発病後の期間が短い者が望ましい。発生頻度の少ない疾患では、やむを得ず発病後数年経過した患者も研究対象に含めざるを得ないことがある。その場合、重症例で発病後間もなく死亡する者は研究対象から除外されることになり、患者は軽症例に偏った選択となる。

3)健康労働者効果

例えば、会社従業員や公務員などの職域集団と一般住民の健康状態(罹患、死亡、症状など)を比較して、職域集団の方が健康状態が良いという結果が出た場合、職域集団では採用時に健康状態の悪いものが排除されている可能性を考慮しなければならない。より健康な人々からなる集団なので、健康労働者効果という。

4)志願者バイアス(回答者バイアス非回答者バイアス)

疫学調査に自発的に参加する人、質問調査に回答する人などは、健康に対する意識が高く、積極的に健康増進に努めている人が多い。また喫煙者も少なく、非参加者、非回答者に比べて、系統的な偏りがあるかもしれない。

5)追跡不能バイアス

コホート研究や介入研究で追跡不能者の頻度が多い場合に起こりうるバイアスである。追跡不能の出現が観察開始時の危険因子の有無や、健康異常の発生と関係がある場合にバイアスが起こりうる。追跡不能によるバイアスを少なくするには、経験的に脱落を20%以下に抑える努力をすべきである。

6)自己選択バイアス

患者および対照を選択するとき、曝露のある者の協力が曝露のない者に比べて悪い、またはその逆の場合があると、研究対象に偏りが生じ、曝露によるリスクが正しく評価されないことがある。

7)さらけ出しバイアス

曝露を受けた者は受けない者に比べて、医療機関受診、健康診断を受ける機会が多いと、非曝露者に比べて早期患者が発見されやすくなり、見かけ上高い罹患率が得られる。このバイアスは後述の情報バイアスでもある。

(2)情報バイアス

観察集団から情報を得るときに起きるバイアスで、測定の違い、誤回答、記憶の違いなど様々な原因が考えられる。

1)思い出しバイアス

症例対照研究で、症例は自分の病気についてよく考えるために、対照よりも良く過去の曝露を覚えているかもしれない。その他いろいろな条件によって思い出しの程度が異なるために起きうるバイアスである。

2)面接者バイアス

ある曝露とあの病気の発生と関係があると考える面接者は患者群についてより熱心に曝露の事実を聞き出そうとするかもしれない。そのようなときに起こりうるバイアスで、曝露の効果が多めに見積もることになる。

3)曝露容疑バイアス

患者と診断した医師が、その病気の発生要因を持っているに違いないと考え、対照に対する場合と異なる熱心さを持って曝露状況を調べたとする。このように予断をもって調査をすると、患者の曝露が多めに見積もられる。

4)診断容疑バイアス

曝露があるか否かが診断に影響を及ぼすことがある。例えば、アスベスト曝露があるものについては、胸部 X 線診断でわずかな変化でも異常ありと判断し、曝露のないものでは同じような変化があっても異常なしと判断するかもしれない。

4. 交絡因子

症例対照研究やコホート研究で健康異常の危険因子を探そうとするときに、調べようとする危険因子以外の表面に出ない別の因子が健康異常の発生に影響を及ぼす場合がある。

例えば、飲酒と食道がんの発生との関係を調べたところ、飲酒者は非飲酒者に比べて食道がんの発生が多いという結果が出たと仮定する。その場合、単純に飲酒が食道がんの発生要因と結論づけることはできない。もし喫煙者の割合が非飲酒者に比べて飲酒者に多く、喫煙が食道がんの発生要因だとすると、やはり同じ結果が得られるはずである。このように表面にでない要因で、健康異常の発生に影響を及ぼす要因を交絡因子という。

5. バイアスや交絡因子に対する対策

(1)研究計画の時点で考慮する方法

1)無作為化

研究対象となる母集団から研究対象者を無作為に選出することによって、系統的な偏りを除外することができ、母集団を代表する標本が得られる。

2)対象の制限

すでにわかっている交絡因子の影響を除くために、非喫煙者のみを研究対象にするというように研究対象者を特定の属性を持つものに限る。

3)マッチング

症例対照研究における症例群と対照群、コホート研究における曝露群と非曝露群を研究対象として選定する場合に、性や年齢分布など除外しようとする交絡因子の影響を除くために、その因子の分布が両群同じくなるようにする方法。

(2)解析の段階で考慮する方法

1)層化

研究対象を交絡因子の値によって層に分けて、それぞれの層ごとに比較する方法。例えば、対象集団を年齢の各層において比較すれば、年齢の影響は除くことができる。

2)標準化

交絡因子の分布が異なる集団を比較する場合、交絡因子の影響を除くために両群の分布が同じであると仮定して比較する方法であり、死亡率や有病率の地域差を比較する場合にしばしば適用される方法。年齢調整死亡率や標準化死亡比はその例である。

3)モデリング

対象とする交絡因子の数が多い場合(例えば、年齢、血圧、肥満の3つの因子の影響を除いた上で、喫煙と虚血性心疾患の罹患との関係を明らかにしたいとき)、層化や標準化による評価が困難であり、これらの因子を数学モデル(多変量)解析に組み込んで、それぞれの因子の影響を明らかにしようとする方法である。

多重ロジスティックモデル、比例ハザードモデルなどがある。

(参考 柳川洋 誤差とその制御、シリーズ「地域医療のための疫学手法」)

調査の妥当性を脅かす要因(武藤 1998)

妥当性の種類	妥当性を脅かす要因		制御方法
外的妥当性 ¹⁾	バイアス	抽出バイアス	無作為抽出
	偶然		統計的検討
内的妥当性 ²⁾	バイアス	選択バイアス	無作為割付
		測定バイアス	測定条件の標準化
		交絡バイアス	無作為割付、層化、標準化
	偶然	平均への回帰	複数回測定、対照群設定
	反応効果		マスク化
	対象の変化	成熟	
脱落			

1): 研究結果を他の事例に一般化できるかどうかを表す概念。

2): 研究結果そのものに信頼性があるかどうかを表す概念。

<統計的仮説の検定>

1. 分布の中心を表すもの(代表値)

①意味が明瞭 ②計算が簡単 ③数学的発展性 ④統計的推論に有用

(1)平均(mean)

μ (母平均)

\bar{x} (標本平均)

相加(算術)平均(arithmetical mean)ともよばれ、最も多用される代表値。

データの数値(変量の値)の合計($\Sigma x = x_1 + x_2 + \dots + x_n$)をデータの個数(n)で割ることにより求められる。すべてのデータの数値を利用するため、極端な数値の影響を受けやすい。平均値と呼ばれるものには、これ以外に相乗(幾何)平均、調和平均などがあるがいずれもあまり一般的ではない。

(2)メジアン(median)Md

データを大きさの順に並べたときその中央に位置する数値で中央値ともいう。

算出方法はデータの個数により次の二通りがある。

- ① データ数が奇数($2n + 1$)とき、 $n + 1$ 番目のデータ。
- ② データ数が偶数($2n$)とき、 n 番目と $n + 1$ 番目のデータの平均値。

(3)モード(mode)Mo

データの中に最もよく出現する値。最頻値ともよばれる。度数分布表では最大度数の階級値を示す。この場合、モードは階級の取り方で変化する。また二峰性の分布の場合、高い方の頂点からそれぞれ第1モード、第2モードという。データの1部を利用する。

(4)代表値の選択と利用

分布の形が左右対称に近い釣り鐘型(正規分布)では、平均値、メジアン、モードは互いに近似した値をとる。この場合の代表値には、計算処理が容易で、統計的推論に優れた平均値が使用される。一方、分布の形が非対称型やL字型であったり、データのばらつきが大きい(平均値より標準偏差が大きい)と、平均値の代表性は低下し意味もわかりにくくなる。こうした場合には、変量の順序のみに関係するメジアンの方が分布から外れた数値の影響を受けにくいと、平均値より代表性に優れている。

モードもメジアン同様、極端な数値の影響を受けにくい。規格品の導入を考えるとといった場合には、モードの利用が適している。

2. データの広がりを示すもの(散布度)

(1)範囲(range)

(2)分散(variance)

$$\mu^2 = \frac{1}{n} \Sigma (x_i - \bar{x})^2 \quad (\text{母分散 population v.})$$

母集団の分散には σ^2 が使われるが、標本集団の分散を母分散の推定値に用いるとき、そのままでは標本分散は母分散の $(n - 1) / n$ 倍になるため、標本集団には分母の n を $n - 1$ に代えた S^2 を用いる。

$$S^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2 \quad (\text{不偏分散 sample v. / 標本分散 unbiased v.})$$

(3)標準偏差(S.D.: standard deviation)

$$\sigma = \sqrt{\frac{1}{n} \Sigma (x_i - \bar{x})^2}$$

分散と同様、標本集団には分母を $n - 1$ に代えた標本標準偏差 S を用いる。

(4)標準誤差(S.E.M.: standard error of mean)

「標本調査」における平均値の信頼度を判断する指標で、標本数が多くなると標準誤差は小さくなる(信頼度は増す)。 σ が不明であれば代わりに S を用いる。

$$S.E. = \frac{\sigma}{\sqrt{n}}$$

平均 μ 、標準偏差 σ の母集団を仮定し、そこから n 個の標本を抜き出し平均を求める。このとき、標本平均の分布は、どのような母集団についても、平均 μ 、標準偏差 σ / \sqrt{n} の正規分布になることが証明されている(中心極限定理)。

標準誤差は、この標本平均のバラツキの程度を示す。

(5)変動係数(C.V.: coefficient of variation)

平均値に対する標準偏差の割合を示す係数で、平均値の異なるデータ間の散布度を比較するのに用いる。

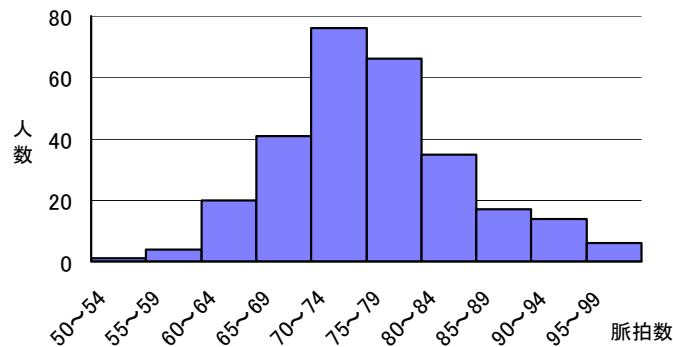
例えば、平成 7 年度の男子の身長(平均±標準偏差)は、幼稚園児(5 歳)111.0 ± 4.70cm、高校生(17 歳)170.8 ± 5.64cm である。両者の変動係数 4.2 %と 3.3 %から、幼稚園児の方が身長に関して個体差が大きい。

$$C.V. = \frac{\sigma}{x} \times 100 (\%)$$

3. 測定データの確率分布

(1)正規分布の性質

次の図はある学校に通う学生 280 名の心拍数の度数分布を表したものである。これを相対度数分布に変換し、全体の面積が1となるように階級の幅を使って縦軸のスケールを変える。そして、データ数を増やしながら階級の幅を小さくしていくと、次第にスムーズで釣り鐘状の「確率密度関数」に近づいていく。度数分布が標本集団である実測値の分布を示すのに対して、確率密度関数は予想される母集団の分布の確率を理論的に示したもので、その曲線下の面積は1となる。正規分布(normal distribution)は連続変量の分布を表す代表的な確率密度関数であり、心拍数以外にも身長や血圧など生体の測定値や測定誤差など多くの変量が正規分布に適合するといわれている。



正規分布は両側に無限に広がる一峰性の分布で、左右対称なため平均値、メジアン、モードはいずれも一致する。正規分布の形は平均値 μ と分散 σ^2 (標準偏差 σ) により完全に決まることから、記号を用いて $N(\mu, \sigma^2)$ と表記している。正規分布と標準偏差の関係には、 μ を中心にした $\pm 1\sigma$ の範囲にデータ全体の 68.3%、 $\pm 2\sigma$ の範囲では 95.5% を含むというたいへん重要な性質がある。よく臨床検査などで集団の平均値を中心に標準偏差の2倍以内を正常域として設定することがあるが、それもこの性質を利用しているといえる。

データが正規分布するかどうかは、ヒストグラムでも見当はつくが、累積相対度数を求め正規確率紙にプロットすることで容易に判定することができる。

(2)正規分布の利用

一般の正規分布 $N(\mu, \sigma^2)$ は平均値と標準偏差の値の違いによってさまざまな形になる。そのうちで平均値0、標準偏差1の正規分布を標準正規分布 $N(0, 1^2)$ とよんでいる。これは一般の正規分布の平均値を $-\mu$ だけ移動し、標準偏差が1になるように横軸のスケールを変えたものであるから、一般の正規分布の値 x は次の式で標準正規分布の値 Z に変換、すなわちデータを標準化することができる。

$$Z = \frac{x - \mu}{\sigma}$$

標準正規分布の横軸上の値0と Z で区切られた曲線下の面積について、いろいろな Z の値について計算した結果をまとめたものを正規分布表という。正規分布は左右対称なので、正規分布表では $Z=\infty$ のときの面積は 0.5 というように片側の数値が示されている。データの標準化により、この正規分布表は一般の正規分布に応用することができる。

(3)その他の理論分布

保健医療領域で扱うほとんどの場合では母集団の母平均 μ と母標準偏差 σ は未知である。しかし母集団が正規分布しなくても、標本集団が十分大きい(データ数 30 以上)ときには、標本平均 \bar{x} とその標準偏差(前述の標準誤差)を母集団の推定値として代用することができる。このとき値 Z は標準正規分布をする。

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

これを利用すれば、標本平均からある確率のもとで母平均の存在する範囲を推測することができる。これを区間推定とよび、確率としては 95%がよく用いられる。この場合、正規分布は左右対称なので、面積が 0.45 になるときの $Z(1.96)$ を用いる。母標準偏差 σ が不明であっても、データ数が多ければ標本標準偏差 S が代用できる。

$$\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}$$

これを平均値の 95%信頼区間という。

しかし、データ数が少ない場合はどうだろうか。小数データを正規分布に当てはめると誤差が大きいことから、母標準偏差 σ に代わって標本標準偏差 S を用いて標準化した値 t が利用される。

$$t = \frac{\bar{x} - \mu}{S / \sqrt{n}}$$

この値 t は自由度 $n - 1$ の t 分布(t-distribution)に従う。

t 分布は左右対称で山型の分布であるが、標本の大きさによって分布の形が変わり、自由度が大きくなるにつれて標準正規分布に近づいていく。

t 分布のほかに標準正規分布から派生した分布に χ^2 分布と F 分布がある。

【演習3】血液データ(添付資料)からヒストグラムを作成しなさい。

4. 仮説検定の key word

(1) 帰無仮説 (null hypothesis : H_0)

仮説検定の原理は背理法である。棄てられる運命にある仮説を帰無仮説という。一方、 $A = B$ という帰無仮説に対して $A \neq B(H_1)$ 、 $A > B(H_2)$ 、 $A < B(H_3)$ という 3 種類の仮説を対立仮説という。統計的仮説には帰無仮説を用いる。

(2) 有意水準 (level of significance : α)

判断の基準となる確率(5%または 1%)を有意水準といい、 α で表わす。帰無仮説の成立する可能性が α 以下のとき、帰無仮説を棄却する。つまり 5%の有意水準という表現は、本当は正しい $H_0(A = B)$ を否定して、誤った H_1 、 H_2 、 H_3 を採用する可能性が 20 回に 1 回あることを意味している。換言すれば、 $P < 0.05$ という表現は、帰無仮説の成立する確率が 0.05 より小さいことを示す。

(3) 第 1 種のエラー (type 1 error : α) と 第 2 種のエラー (type 2 error : β)

検定に際して、正しい H_0 を誤って棄てる誤りを第 1 種のエラーといい、 H_0 が誤っているにもかかわらず、棄却しない誤りを第 2 種のエラーという。この両者は一方を小さくすると他方は大きくなる関係にあり、バランスを取る必要から、 α は 0.01 ~ 0.05 に設定される。

(4) 両側検定 (two tailed test) と 片側検定 (one tailed test)

母平均 μ_1 、 μ_2 をもつ 2 つの母集団があり、 μ_1 と μ_2 の差を検定する時の H_0 を $\mu_1 = \mu_2$ とする。このときの対立仮説は $\mu_1 > \mu_2$ または $\mu_1 < \mu_2$ である。仮説が正しいにもかかわらず、 $\mu_1 > \mu_2$ と結論を出す確率を α_1 、 $\mu_1 < \mu_2$ と結論を出す確率を α_2 とすると、 $\mu_1 = \mu_2$ 以外の結論を出す確率は $(\alpha_1 + \alpha_2)$ である。このように $(\alpha_1 + \alpha_2)$ のような確率水準を採用する場合を両側検定という。正規分布や t 分布を応用して検定する場合、 $\alpha = \alpha_1 + \alpha_2$ であるから、 $\alpha = 0.025$ に相当する Z 値や t 値を用いる。

一方、明らかに $\mu_1 \geq \mu_2$ である場合には、誤って $\mu_1 < \mu_2$ となる確率をゼロとして有意水準 $\alpha = \alpha_1$ を採用する。これが片側検定である。

<データの統計処理>

1. 平均値の差の検定

正規分布 母分散が既知のとき(大標本の場合 σ を S で代用)

t 分布 母分散が未知のとき(小標本の場合)

- (1)独立 2 標本の平均の差の検定(母分散既知か大標本)
- (2)独立(対応のない) 2 標本の平均の差の検定 (母分散未知)
 - 母分散が等しいとき (Student's t -test)
 - 母分散が等しくないとき (Welch の方法)
- (3)対応のある 2 群の小標本平均の差の検定 (paired t -test)

【例題2】次の表は、昨年同時期にある歯科医院を初めて受診し、その後 1 年ぶりに来院にした幼児 8 名 (A) と、年間を通じて定期的に指導を受けてきた幼児 8 名 (B) の年間う蝕発生歯面数を示している。両者の平均はそれぞれ 5.64 と 3.14 であった。

この結果から、この間の定期的な受診によるう蝕抑制効果があったと判断してよいか。

幼児 A	5	3	8	6	0	5	4	3
幼児 B	4	2	0	0	3	2	5	3

(有効数字 3 桁、危険率 5%)

(解説)

等質の母集団から抽出された独立 2 標本 (n_A, n_B) において、標本平均の差 ($\bar{x}_A - \bar{x}_B$) は自由度 ($n_A + n_B - 2$) の t 分布となることを利用する

	幼児 A	幼児 B
標本数	n_A	n_B
標本平均	\bar{x}_A	\bar{x}_B
不偏分散	S_A^2	S_B^2

$$n_A = 8$$

$$\bar{x}_A = 1 / 8 (5 + 3 + \dots + 3) = 4.25$$

$$S_A^2 = 1 / (8 - 1) \{ (5 - 4.25)^2 + (3 - 4.25)^2 + \dots + (3 - 4.25)^2 \} = 5.64$$

$$n_B = 8$$

$$\bar{x}_B = 1 / 8 (4 + 2 + \dots + 3) = 2.38$$

$$S_B^2 = 1 / (8 - 1) \{ (4 - 2.38)^2 + (2 - 2.38)^2 + \dots + (3 - 2.38)^2 \} = 3.13$$

(1) H_0 の設定 (両者の平均年間う蝕発生歯面数に差はない)

(2) 母分散が等質かどうかの確認 (F 検定)

① 両者の母分散に差はない (H_0')

② F_0 値の計算 ($F_0 > 1$)

$$S_A^2 > S_B^2 \text{ より } F_0 = S_A^2 / S_B^2 = 5.64 / 3.13 = 1.80$$

③ F 値を求める

$$F (n_A - 1, n_B - 1, \alpha) = F (7, 7, 0.05) = 3.79 \quad \rightarrow \text{自由度: 分子(横) 分母(縦)}$$

④ $F_0 < F$ より H_0' は棄却できない \therefore 母分散は等しい

(3) 共分散 ω^2 を求める

$$\begin{aligned} \omega^2 &= \frac{\sum (x_{Ai} - \bar{x}_A)^2 + \sum (x_{Bi} - \bar{x}_B)^2}{n_A + n_B - 2} \\ &= \frac{(n_A - 1) S_A^2 + (n_B - 1) S_B^2}{n_A + n_B - 2} \\ &= \frac{(8 - 1) \times 5.64 + (8 - 1) \times 3.13}{8 + 8 - 2} = 4.39 \end{aligned}$$

(4) t_0 値を求め t 値と比較する

$$t_0 = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{\omega^2 (1/n_A + 1/n_B)}} = \frac{|4.25 - 2.38|}{\sqrt{4.39 (1/8 + 1/8)}} = 1.79$$

$$t (n_A + n_B - 2, \alpha) = t (14, 0.05) = 2.145$$

$t_0 < t$ より H_0 は棄却できない

(5) 結論 (危険率 5% で両者の年間う蝕発生歯面数に差はあるとはいえない)

【演習4】を行いなさい。(データはエクセルファイルを参照)

2. 3群以上のデータの取り扱い(分散分析: ANOVA analysis of variance)

標本分散の違いを検定する

(1)一元配置分散分析(one way ANOVA) 1因子により分類される多群

(2)二元配置分散分析(two way ANOVA) 2因子により分類される多群

繰り返しのある二元配置

繰り返しのない二元配置

一元配置分散分析について例題で示す

【例題3】抜歯後疼痛を押さえるため患者に3種類の鎮痛剤を経口投与した。作用の発現までの時間(分)は次のようであった。薬剤により差があるといえるか。

(解説)

H_0 : 薬剤による作用発現時間(分)に差はない。

薬剤 A 投与群	薬剤 B 投与群	薬剤 C 投与群
185	200	265
225	186	214
215	178	226
245	196	224
232		248

A 群の平均 220.4 B 群の平均 190 C 群の平均 235.4 全体の平均 217.0714286

1) 群間の変動を求める

群間の偏差平方和 4666.528571

2) 群内の変動を求める

A の偏差平方和 2043.2

B の偏差平方和 296

C の偏差平方和 1711.2

群内の偏差平方和 4050.4

全体(群間+群内)の偏差平方和 8716.928571

3) 分散分析表にまとめる

変動要因	偏差平方和	自由度	分散	分散比	P
群間変動	4666.528571	2	2333.264286	6.33663518	0.0148
群内変動	4050.4	11	368.2181818		
総変動	8716.928571	13			

4) 判定

危険率 5%で薬剤による作用発現時間(分)に差がある。

(分散分析ではどの薬剤間に差があるかは分からない。薬剤間の差は多重比較で行うが省略する。)

【演習5】を行いなさい。(データはエクセルファイルを参照)

<対になったデータの取り扱い(相関と回帰)>

1. 相関係数 (r : correlation coefficient)

相関図(散布図)に分布する点が、回帰直線の近辺に集中して分布していれば、 x の任意の値に対する y の中心的な値を推定することができる。相関係数はデータの分布が回帰直線の回りに集中する程度を表す。

すべてのデータは (x_i, y_i) という対のデータとして書き表される。そこで、 X 、 Y 成分の変動として各々の偏差平方和(T)を考える。

$$X \text{ 成分の変動 } T_x = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$Y \text{ 成分の変動 } T_y = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

X 、 Y 成分間に関連する変動を両者の偏差の積で考える。それを共変動という。

$$X \text{ と } Y \text{ の共変動 } T_{x,y} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

この共変動は、2変数の関係が直線的であるほど正または負の大きな値になるため、 X 、 Y 成分の関連の強さを示す指標となる。

X 、 Y 成分の共変動を各々の変動の平方根の積で除したものを相関係数(r)という。

$$r = \frac{T_{x,y}}{\sqrt{T_x T_y}}$$

相関係数の有意性の検定

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

t 値が、自由度 $n-2$ の t 分布することを利用して、 t 分布表から判定する

2. 直線回帰 (linear regression)

変数 X 、 Y に相関関係があるとき、その分布の中心的傾向を予測する直線を回帰直線という。

求める直線の方程式を $y = ax + b$ と仮定し、 $P(x_i, y_i)$ から垂線をおろし、直線との交点を $Q(x_i, y')$ とすれば、 PQ 間の差は (y_i, y') となる。

変数 X から Y を予測する場合の誤差を最小とするのは $\sum (y_i, y') = 0$ であり、 $\sum (y_i, y')^2$ が最小となる a 、 b を求めればよい。

$PQ^2 = \sum (y_i, y')^2 = \sum \{y_i - (ax_i + b)\}^2$ を最小とする a 、 b は

$$a = \frac{T_{x,y}}{T_x}$$

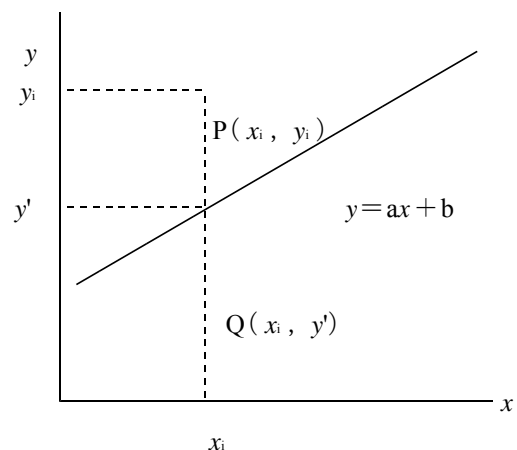
$$b = \bar{y} - a \bar{x}$$

となる。

$y = ax + b$ に a 、 b を代入すると

$$y - \bar{y} = \frac{T_{x,y}}{T_x} (x_i - \bar{x})$$

これを X から Y への回帰直線という。



【演習6】を行いなさい。(データはエクセルファイルを参照)

(付)参考となるサイトの URL

(参考・引用資料は次回のプリントで紹介する。)

Web 独習教材「ハンバーガーショップで学ぶ楽しい統計学」

<http://kogolab.chillout.jp/elearn/hamburger/index.html>

エクセルの利用法

<http://www1.tcue.ac.jp/home1/abek/htdocs/stat/index.html>

平成 28 年度 実験計画法

第3回(ノンパラメトリック検定)

4月 18日(月) 10:25 ~ 11:55

場所:楠元 4号館3階 パソコン室

担当:加藤一夫

<離散量(discrete variate)データの取り扱い(χ^2 検定)>

構成集団中である特性を持ったものの割合が問題となることが多い。
分割表を利用する。

1. 適合度の検定(1 要因: 単純集計の場合)

1 つの母集団の性質を検討する。

観測値 O 、期待値 E とするとき、 χ_0^2 値より一致度を算出する。

$$\chi_0^2 = \sum \frac{(O-E)^2}{E}$$

一般には連続性の補正(Yate's の補正)を行う。

$$\chi_0^2 = \sum \frac{(|O-E| - 0.5)^2}{E}$$

この χ_0^2 値を自由度 $n-1$ の χ^2 値と比較し、 $\chi_0^2 \geq \chi^2(\alpha)$ であれば、危険率 α で H_0 を棄却する。

2. 独立性の検定(2 要因: $l \times m$ 分割表の場合)

2 つの母集団の性質を比較する。

次の式は 2×2 分割表を利用した場合のみ利用できる。

$$\chi_0^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (n=a+b+c+d, a, b, c, d \neq 0)$$

一般には連続性の補正(Yate's の補正)を行う。

$$\chi_0^2 = \frac{n(|ad-bc| - n/2)^2}{(a+b)(c+d)(a+c)(b+d)}$$

この χ_0^2 値を自由度 $(l-1) \times (m-1)$ の χ^2 値と比較し、 $\chi_0^2 \geq \chi^2(\alpha)$ であれば、危険率 α で H_0 を棄却する。

【例題4】小学生 80 名について、おやつのお味嗜好とう蝕発生の有無について調査した。両者は関連するか。(有意水準 5%)

(解説)

H_0 : おやつのお味嗜好とう蝕発生の有無に関連はない。

	う蝕あり	う蝕なし	計
嗜好あり	30	20	50
嗜好なし	8	22	30
計	38	42	80

$$\chi_0^2 = \frac{80 \times (|30 \times 22 - 8 \times 20| - 80/2)^2}{38 \times 42 \times 30 \times 50} = 7.07$$

$$\chi^2(1, 0.05) = 3.841$$

$\chi_0^2 > \chi^2$ より H_0 は棄却される。

危険率 5% で有意な差がある。

【演習7】、【演習8】を行いなさい。(データはエクセルファイルを参照)

3. McNemar の検定 (対応のある2標本の比率の差の検定)

同一被験者を対象に2度同じ検査をするといった、ある事象前後での変化が有意かどうか (H_0 :前後の変化は有意ではない。) を検定するもの。

次の表で、1回目と2回目で反応の異なる b と c に注目する。

		1回目の反応		計
		+	-	
2回目 の反応	+	a	b	a + b
	-	c	d	c + d
計		a + c	b + d	

a (1回目、2回目ともに+) と d (1回目、3回目ともに-) は情報にならない。

変化のなかった場合の期待値 $E = \frac{b + c}{2}$ と観察値 b, c の差を利用する。

$$\chi^2 = \frac{(b - \frac{b+c}{2})^2}{\frac{b+c}{2}} + \frac{(c - \frac{b+c}{2})^2}{\frac{b+c}{2}} = \frac{(b - c)^2}{b + c}$$

連続性の補正をする場合は

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

この χ^2 値を自由度 1 の χ^2 値と比較し、 $\chi^2 \geq \chi^2(\alpha)$ であれば、危険率 α で H_0 を棄却する。

【演習 9】偏頭痛で悩む 30 名の被験者に、新しく開発された鎮痛剤とそのプラセボをランダムに 1 回ずつ投与しその効果を判定した。効果があった場合を「+」、無かった場合を「-」として次の表にまとめた。

この鎮痛剤に偏頭痛を抑える効果があるといえるか。

	鎮痛剤	プラセボ
1	+	+
2	-	-
3	+	-
4	+	-
5	-	+
6	-	-
7	+	+
8	+	-
9	-	-
10	+	+
11	-	-
12	+	+
13	+	-
14	-	-
15	+	+
16	+	-
17	+	-
18	+	-
19	+	-
20	+	+
21	-	+
22	+	-
23	+	+
24	+	+
25	-	-
26	-	-
27	+	+
28	+	-
29	-	+
30	-	-

<順位で示されるデータの取り扱い(ノンパラメトリック検定)>

1. Wilcoxon の T 検定 対応のある 2 群の比較

例題で示す

【例題5】患者 10 名について右側と左側で臼歯部の咬合圧に差があるか。

H₀ : 左右の咬合圧に差はない。 (kg)

患者	1	2	3	4	5	6	7	8	9	10
左側	49	35	40	49	34	42	52	41	32	40
右側	53	42	46	46	39	42	49	44	36	49
差	-4	-7	-6	3	-5	0	3	-3	-4	-9

1) 符合を無視して並び替える。(同順位 0 は除外)

3 3 -3 -4 -4 -5 -6 -7 -9

2) 符合を無視して順位をつける。

2 2 2 4.5 4.5 6 7 8 9

3) 符合つき順位を求める。

2 2 -2 -4.5 -4.5 -6 -7 -8 -9

左側 T₊ = 2+2 = 4

右側 T₋ = 2+4.5+4.5+6+7+8+9 = 41

4) 小さい方の T₊ 値を T 表 (n=9 のとき) の T 値と比較し、

T₊ 値がそれよりも小さければ仮説を棄却する。

(確率の求め方)

n ≤ 25 のとき Wilcoxon 検定表を利用。

n > 25 のとき

$$\mu_T = \frac{n(n+1)}{4} \quad \text{と}$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{4}} \quad \text{から}$$

$$Z = \frac{T - \mu_T}{\sigma_T} \quad \text{を算出する。}$$

正規分布表で Z を求める。

Wilcoxon 検定表 : T の有意点

片側確率	P<0.025	P<0.005
両側確率	P<0.05	P<0.01
n = 6	0	
7	2	
8	3	0
9	5	1
10	8	3
11	10	5
12	13	7
13	17	9
14	21	12
15	25	15
16	29	19
17	34	23
18	40	27
19	46	32
20	52	37
21	58	42
22	65	48
23	73	54
24	81	61
25	89	68
26	98	75
27	107	83
28	116	91
29	126	100
30	137	109
31	147	118
32	159	128
33	170	138
34	182	148
35	195	159
36	208	171
37	221	182
38	235	194
39	249	207
40	264	220
41	279	233
42	294	247
43	310	261
44	327	276
45	343	291
46	361	307
47	378	322
48	396	339
49	415	355
50	434	373

2. Mann-Whitney の U 検定 独立した 2 群の比較

例題で示す

【例題 6】妊娠中の女性患者 7 名と同世代の女性患者(非妊娠) 13 名の唾液のグルコースクリアランス値を測定しデータを得た。両群には差があるか。

H_0 : グルコースクリアランス値に差はない。(2 群の順序関係に偏りはない)

	1	2	3	4	5	6	7	8	9	10	11	12	13	
女性(非妊娠)	400	270	220	190	190	180	170	170	170	160	130	110	80	
妊娠女性	470	380	360	290	280	220	170							
	①	②	③	④	⑤	⑥	⑦							(mg/dl)

1) U 値を求める(ここでは妊娠女性に注目し、それを n_1 とする)。

個々の点に注目して、その点より大きい他方の群のデータ数を調べ合計する。

① ② ③ ④ ⑤ ⑥ ⑦ 注 ⑥は 2 と 3 の平均

順位 0 1 1 1 1 2.5 7.5 ⑦は 6 と 9 の平均

$U \text{ 値} = 0 + 1 + 1 + 1 + 1 + 2.5 + 7.5 = 14$

2) この U 値を Mann-Whitney の U 表で $n_1 = 7$ のときの値と比較し、それよりも小さければ仮説を棄却する。(U 値が小さいほど群間の差は大きくなる)

(確率の求め方)

$n_1 \leq 20$ かつ $n_2 \leq 20$ のとき Mann-Whitney 検定表を利用。

n_1, n_2 の一方が 20 以上するとき

$$\mu_U = \frac{n_1 n_2}{2} \text{ と } \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \text{ から } Z = \frac{U - \mu_U}{\sigma_U} \text{ を算出。}$$

正規分布表で Z を求める。

Mann-Whitney 検定表 : U 値の有意点(両側確率)

$P < 0.05$

n2=	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	=n2
n1=1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	n1=1
2	-	-	-	-	0	0	0	0	1	1	1	1	1	2	2	2	2	2
3	-	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	3
4	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13	4
5		2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20	5
6			5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	6
7				8	10	12	14	16	18	20	22	24	26	28	30	32	34	7
8					13	15	17	19	22	24	26	29	31	34	36	38	41	8
9						17	20	23	26	28	31	34	37	39	42	45	48	9
10							23	26	29	33	36	39	42	45	48	52	55	10
11								30	33	37	40	44	47	51	55	58	62	11
12									37	41	45	49	53	57	61	65	69	12
13										45	50	54	59	63	67	72	76	13
14											55	59	64	67	74	78	83	14
15												64	70	75	80	85	90	15
16													75	81	86	92	98	16
17														87	93	99	105	17
18															99	106	112	18
19																113	119	19
20																	127	20

3. Spearman の順位相関

例題で示す。

【例題 7】歯周炎患者の唾液の潜血状態を検査紙とヘモグロビンの定量法で測定した。両者の相関を求めよ。

H_0 : 両者に相関はない。(無相関検定)

検査紙法 A	0	0	0	1	1	2	2	3	3	3
定量法 B	0.01	0.02	0.04	0.03	0.08	0.08	0.14	0.12	0.30	0.85

1) r_s 値を求める。

A の順位	2	2	2	4.5	4.5	6.5	6.5	9	9	9
B の順位	1	2	4	3	5.5	5.5	8	7	9	10
順位差	1	0	-2	1.5	-1	1	-1.5	2	0	-1
d^2	1	0	4	2.25	1	1	2.25	4	0	1

$$n = 10, \quad \Sigma d^2 = 16.5, \quad r_s = 1 - \frac{6 \times 16.5}{10^3 - 10} = 0.9$$

2) この r_s 値を Spearman の順位相関 ($n = 10$ のとき) の値と比較し、それよりも大きければ仮説を棄却する。

(確率の求め方)

$n \leq 30$ のとき Spearman 検定表を利用。

$n > 30$ のとき

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}} \quad \text{が、}$$

自由度 $n - 2$ の t 分布することを利用して、 t 分布表から判定する

Spearman の順位相関係数
検定表 : r_s の有意点

片側確率	$P < 0.025$	$P < 0.005$
両側確率	$P < 0.05$	$P < 0.01$
$n = 5$	1.000	-
6	0.886	1.000
7	0.786	0.929
8	0.738	0.881
9	0.700	0.833
10	0.648	0.794
11	0.618	0.755
12	0.587	0.727
13	0.560	0.703
14	0.539	0.675
15	0.521	0.654
16	0.503	0.635
17	0.485	0.615
18	0.472	0.6
19	0.460	0.584
20	0.447	0.57
21	0.435	0.556
22	0.425	0.544
23	0.415	0.532
24	0.406	0.21
25	0.398	0.511
26	0.390	0.501
27	0.382	0.491
28	0.375	0.483
29	0.368	0.475
30	0.362	0.467

参考資料

バイオサイエンスの統計学 南江堂 1990

(使用した検定表は、ここから引用した。)

臨床医のための医学統計 メディカルコア 1983

医学保健学の例題による統計学 現代数学社 1982

生物統計学入門 培風館 1975

ノンパラメトリック法 日本文化科学社 1964

追加資料

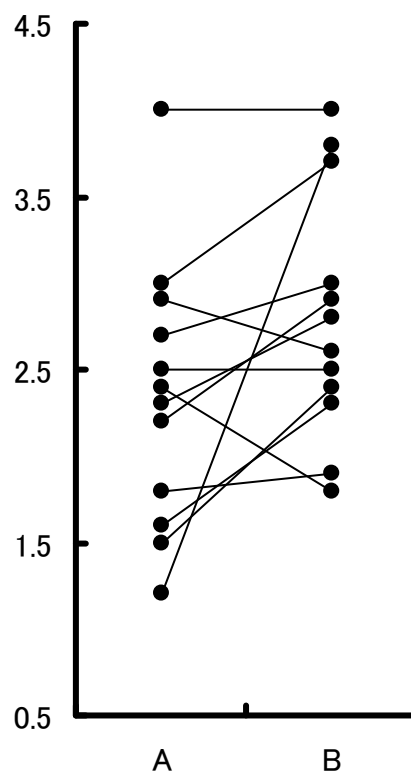
1. パラメトリック検定とノンパラメトリック検定*

検 定	パラメトリック検定	ノンパラメトリック検定*
独立(対応のない) 2 標本の差の検定	t-test Z 検定 F 検定	Mann-Whitney の U 検定(順位和検定)
対応のある 2 標本の差の検定	paired t-test	Wilcoxon の T 検定(順位付き符号検定) 符号検定 McNemar 検定
独立した多群の差の検定	one way ANOVA	Kruskal-Wallis 検定
対応のある多群の差の検定	two way ANOVA	Freedman 検定
相 関	Peason の相関係数	Speaman の順位相関 Kendall の順位相関
分布の適合度		χ^2 検定

*: 母集団の特性を規定する値について何ら特殊な仮説を含まない統計推理法(岩原)

2. ノンパラメトリック検定の演習

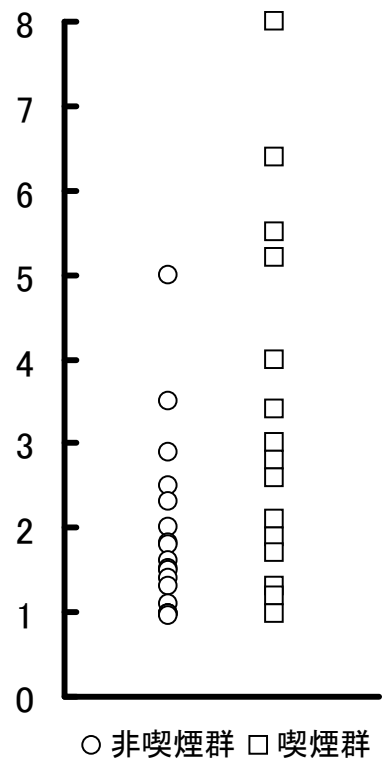
- ① 12 人の患者に、A , B2 種の利尿剤を日を変えて投与し、その効果(尿量)を比較したところ下記データを得た。
両利尿剤の効果に差はあるといえるか。



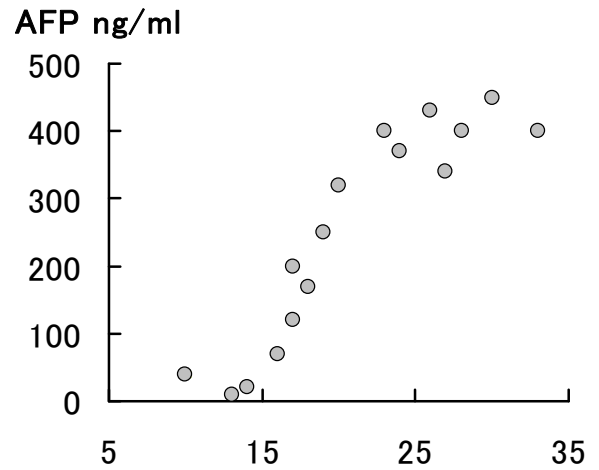
	a	b	c	d	e	f	g	h	I	j	k	l
A	1.8	2.5	2.7	2.9	1.2	4.0	2.4	1.6	1.5	2.3	2.2	3.0
B	1.9	2.5	3.0	2.6	3.8	4.0	1.8	2.3	2.4	2.8	2.9	3.7
d	0.1	0.0	0.3	-0.3	2.6	0.0	-0.6	0.7	0.9	0.5	0.7	0.7

② 老人の血中 CEA 値を、喫煙者と非喫煙者で比較したところ、次のデータを得た。両者に差があるといえるか。

非喫煙群	喫煙群
5	8
3.5	6.4
2.9	5.5
2.5	5.2
2.3	4
2	3.4
1.8	3
1.8	2.8
1.6	2.6
1.5	2.1
1.5	1.9
1.4	1.7
1.3	1.3
1.1	1.2
<1.0	<1.0
<1.0	
<1.0	



③ 妊婦血中の胎児性蛋白濃度 (AFP) 値と妊娠週数の間に次の関係が認められた。両者の相関を求めよ。



妊娠週数	10	13	14	16	17	17	18	19	20	23	24	26	27	28	30	33
AFP ng/ml	40	10	20	70	120	200	170	250	320	400	370	430	340	400	450	400

妊娠週数の順位	1	2	3	4	5.5	5.5	7	8	9	10	11	12	13	14	15	16
AFP 値の順位	3	1	2	4	5	7	6	8	9	13	11	15	10	13	16	13
順位差 d_i	-2	1	1	0	0.5	-1.5	1	0	0	-3	0	-3	3	1	-1	3
d_i^2	4	1	1	0	0.25	2.25	1	0	0	9	0	9	9	1	1	9