EZR 基本解析チャレンジシート - Answer

はじめに

このファイルは統計解析のソザイヤサン（https://toukei.sozaiya-san.jp/）において配布している下記ファイルを使って EZR で統計解析の練習をするための練習シートの答えです。

必要なもの

・基本統計チャレンジデータ.rda（EZR 用のデータセット）

その他

・数値は小数第 1 位まで、相関係数 r、P 値は小数第 3 位までとする（※小数点の最終桁は四捨五入により答えが±0.001 分だけ異なる場合があります）。

データセットの概要

○○県××市の老人ホームに入所している高齢者 100 名を対象とした調査である（架空データ）。ベースラインと 1 年後に以下について測定した（MMSE はベースラインのみ）。
ベースラインに測定した変数は「B.」（Before）、 1 年後は「A.」（After）がそれぞれ変数名の頭についている。

調査内容（変数一覧）

変数名	説明・単位
ID	対象者の ID
性別	0：男性, 1：女性
年齢	歳
身長	cm
体重	kg
エネルギー摂取量	kcal
MMSE	0-30 点　認知症の検査 23 点以下は認知症の疑いがある
食行動.傾眠	0：非該当, 1：該当 「食事中に眠ってしまう」
食行動.拒否	0：非該当, 1：該当 「食事を拒否してしまう」
食行動.むせ	0：非該当, 1：該当 「食事中にむせてしまう」
幸福度	Q.自分の幸福度はどの程度だと思いますか？ 1：不幸　2：少し不幸　3：普通　4：少し幸福　5：幸福
食事満足度	Q.現在の食事に満足していますか？ 1：不満　2：少し不満　3：普通　4：少し満足　5：満足

チャレンジ１【C1】Answer
まずは、調査した対象者がどのような人たちか知ろう

01 以下の変数を【アクティブデータセット → 変数の操作 → 計算式を入力して新たな変数を作成する】から作成しなさい。

変数「B.BMI」… B.体重/(B.身長/100)^2

変数「A.BMI」… A.体重/(A.身長/100)^2

変数「認知症」… ifelse(B.MMSE<=23,1,0)

02 調査対象者の様子（背景）を確認するために、男女に分けた背景表（サマリー表）を、【グラフと表 → サンプルのサマリーデータの背景表の出力】から作成し、表を完成させなさい。

	男性	**女性**
	n = 35	n = 65
	mean（SD）	mean（SD）
年齢，歳	85.5(8.2)	84.8(7.0)
身長，cm	165.6(6.3)	153.5(5.5)
体重，kg	62.2(10.8)	49.2(7.5)
エネルギー摂取量，kcal	1404.9(87.2)	1333.7(74.1)
MMSE，点	21.1(5.7)	20.6(6.1)
BMI，kg/m^2	22.6(3.4)	20.9(2.7)
	中央値（四分位範囲）	中央値（四分位範囲）
幸福度，点	3(3-4)	3(2-4)
食事満足度，点	4(3-4)	4(3-5)
	%	%
食事中の傾眠率	34.3	16.9
食事の拒否率	8.6	9.2
食事中のむせ率	54.3	47.7
認知症有病率	57.1	55.4

チャレンジ２【C２】
男性と女性には、どの測定項目に統計的有意差があるか調べよう

01 以下は、【C1】で作成した表である。男性と女性の差を検定し、以下の表の P 値を埋めなさい。尚、割合の差の検定はフィッシャーの正確検定を用いなさい。

対応のない t 検定で検定

	男性	女性	
	n = 35	n = 65	***P***
	mean（SD）	mean（SD）	
年齢，歳	85.5(8.2)	84.8(7.0)	0.675
身長，cm	165.6(6.3)	153.5(5.5)	<0.001
体重，kg	62.2(10.8)	49.2(7.5)	<0.001
エネルギー摂取量，kcal	1404.9(87.2)	1333.7(74.1)	<0.001
MMSE，点	21.1(5.7)	20.6(6.1)	0.707
BMI，kg/m^2	22.6(3.4)	20.9(2.7)	0.006

マンホイットニーU検定で検定

	中央値（四分位範囲）	中央値（四分位範囲）	***P***
幸福度，点	3(3-4)	3(2-4)	0.704
食事満足度，点	4(3-4)	4(3-5)	0.523

フィッシャーの正確検定で検定

	%	%	***P***
食事中の傾眠率	34.3	16.9	0.080
食事の拒否率	8.6	9.2	1.000
食事中のむせ率	54.3	47.7	0.676
認知症有病率	57.1	55.4	1.000

チャレンジ３【C３】
ここまでの結果をまとめて、本格的な統計解析の準備をしよう

01	【C２】で男女比較をした結果をもとに、以下の文章の空欄を埋めなさい。また、正しい文章にチェックしなさい。

男性と女性を比較したところ、有意差が見られた変数は

（　身長　体重　エネルギー摂取量　BMI　）であり、

その他の項目は有意差が見られなかった。有意差が見られた変数を扱う場合、

- ✓男女を分けて検定したほうが良い可能性がある。
- □男女を分けずに検定したほうが良い可能性がある。

次に進む前に、必ず以下をよく読んで理解しよう

ポイント

特に有意差が見られなくても、ｎ数がある程度あれば男女分けて解析を進める方が一般的であり、解釈しやすくなる。ｎ数が少なく、男女を分けることができない場合は、基礎解析の手法では検定できない。高度な解析手法である多重比較検定を使用し、調整を行う必要がある。

なぜ分けるのかをもう少し詳しく

男女を分けずに認知症の有病群と非有病群を比較したところ、「有病群の体重が有意に小さかった。しかし、有病群の方が女性の割合が大きいことも分かった」。もしこのような結果になった場合、認知症の有病により体重が減るのか、女性の方が認知症に罹患しやすく認知症有病群に女性が多く存在し、その結果体重に差が出たのか判断できない。

性別以外にも、年齢も同じことがいえる。有病群と非有病群を比較したところ、有病群の体重が低かったが、同時に年齢も高齢であった。一般的に高齢になるほど体重は低くなる。このように、有病群の年齢が高いから体重が小さいのか、年齢とは独立して、有病群は体重が小さいのか判断できない。

このような問題を解決するためには、男女や年齢階級で分けて検定するか、性別・年齢を調整できる高度な解析手法を用いる必要がある。

02	これから、認知症の有病者と非有病者を比較することで、認知症の有無で何が異なるのか調べていきたい。これを探ることで、認知症の原因解明が期待できる（認知症の有無で差がある項目は認知症の原因の可能性がある）。以下の項目について自分の予想に合う選択肢にチェックを入れなさい。自分の経験や知識をもとに考えてみましょう。

この問題に正解はありません。【C4】から実際に検定していくことで、ここで立てた予想、すなわち研究仮説を検証していきながら、認知症の原因は何かを探っていきます。

チャレンジ４【C４】
認知症の有無で分けた際、標本数はいくつか、年齢はどうかを男女分けて調べてみよう

01 以下について【グラフと表 → サンプルのサマリーデータの背景表の出力】から計算し、文章・表の空欄および選択肢にチェックしなさい。また、P 値は検定を行い計算しなさい。

男性の認知症有病群は（２０）人、非有病群は（１５）人であり、各群の人数が少ないため

□パラメトリック検定を使うほうが良い

□ノンパラメトリック検定を使うほうが良い

女性の認知症有病群は（３６）人、非有病群は（２９）人であり、群の人数が少ないため

□パラメトリック検定を使うほうが良い

□ノンパラメトリック検定を使うほうが良い

１グループの人数が 30 人以上であればパラメトリック検定を使う。女性は微妙なところだが、男性に合わせてノンパラメトリック検定を使うほうが良いだろう。

P 値の計算は、ノンパラメトリック検定のマンホイットニーのＵ検定を使う

ベースラインの**男性**	**認知症有病群**	**非有病群**	
	n = 20	n = 15	***P***
	mean（SD）	mean（SD）	
年齢，歳	84.4(8.0)	86.9(8.4)	0.393

ベースラインの**女性**	**認知症有病群**	**非有病群**	
	n = 36	n = 29	***P***
	mean（SD）	mean（SD）	
年齢，歳	85.4(7.6)	84.1(6.2)	0.269

02 年齢に差があると、年齢を考慮しなければならない。上記の結果から、今回のデータを解析する上で年齢を考慮する必要があるかどうか考えながら、選択肢にチェックしなさい。

男女それぞれ認知症有病群と非有病群の年齢を比較したところ、

□有意な差が見られた

□有意な差が見られなかった

そのため、認知症有病群と非有病群について比較する際に、

□年齢を調整せずに検定していく

□年齢調整や年齢階級別で検定していく必要がある

チャレンジ５【C５】
まずは男性について、認知症有病群と非有病群で何が違うのか検討していこう

01 ベースラインの男性の認知症有病群と非有病群の差を検定し、以下の表を埋めなさい。尚、割合の差の検定はフィッシャーの正確検定を用いなさい。

男性	認知症有病群	非有病群	
	n = 20	n = 15	***P***
	mean（SD）	mean（SD）	
年齢，歳	84.4(8.0)	86.9(8.4)	0.393
身長，cm	165.7(6.2)	165.5(6.6)	0.868
体重，kg	62.3(10.7)	62(11.3)	0.764
エネルギー摂取量，kcal	1417.4(97.5)	1388.3(71)	0.405
MMSE，点	16.9(3.6)	26.6(2)	<0.001
BMI，kg/m^2	22.6(3.3)	22.6(3.5)	0.961
	中央値（四分位範囲）	中央値（四分位範囲）	***P***
幸福度，点	3.00 [2.75 - 4.00]	3.00 [3.00 - 4.00]	0.634
食事満足度，点	4.00 [3.00 - 4.00]	4.00 [4.00 - 5.00]	0.070
	%	%	***P***
食事中の傾眠率	35.0	33.3	1.00
食事の拒否率	10.0	6.7	1.00
食事中のむせ率	55.0	53.3	1.00
認知症有病率	100.0	0.0	―――

チャレンジ6【C6】
次は女性について、認知症有病群と非有病群で何が違うのか検討していこう

01 ベースラインの女性の認知症有病群と非有病群の差を検定し、以下の表を埋めなさい。尚、割合の差の検定はフィッシャーの正確検定を用いなさい。

女性	認知症有病群 n = 36	非有病群 n = 29	*P*
	mean（SD）	mean（SD）	
年齢，歳	85.4(7.6)	84.1(6.2)	0.269
身長，cm	152.6(5.8)	154.6(5)	0.198
体重，kg	49.4(8.7)	49(5.8)	0.571
エネルギー摂取量，kcal	1333.6(81.8)	1333.7(64.8)	0.514
MMSE，点	16.2(4.4)	26.1(2.2)	<0.001
BMI，kg/m^2	21.1(3)	20.5(2.4)	0.500
	中央値（四分位範囲）	中央値（四分位範囲）	*P*
幸福度，点	3.00 [2.00 - 4.00]	3.00 [3.00 - 4.00]	0.408
食事満足度，点	4.00 [3.00 - 4.00]	5.00 [4.00 - 5.00]	<0.001
	%	%	*P*
食事中の傾眠率	25.0	6.9	0.094
食事の拒否率	11.1	6.9	0.684
食事中のむせ率	50.0	44.8	0.804
認知症有病率	100.0	0.0	―――

チャレンジ７【C７】
ここまでの結果をまとめて、【C３】の予想を検討してみよう

01 【C５】と【C６】において有意差が見られた変数をまとめなさい。

認知症有病群で高かった変数

男性	女性

認知症有病群で低かった変数

男性	女性
MMSE	MMSE 食事満足度

02 以下の文章の空欄および選択肢にチェックを入れなさい。【　・　】は正しい文章に〇をつけなさい。丸数字は同じ言葉が入ることを意味する。

MMSE が認知症有病群で低かったのは、

□今回の結果からは理由はわからない。

☑群分けの際に MMSE を用いたため当然である。

【女性】のみ（①食事満足度）が認知症有病群で有意に低い値であった。

このことから女性において、

□認知症の有病には、①が原因であると考えられる。

□認知症が原因となり、結果として①が低下すると考えられる。

☑認知症と①は関係していることが考えられる。

□今回の結果からは認知症と①の関係はわからない。

ここまでの解析は【横断的解析】である。

次に進む前に、必ず以下をよく読んで理解しよう

【C３】で予想した結果と比較してどうだっただろうか？おそらく予想したほど結果は出なかったであろう。しかし、今回は n 数が少なくノンパラメトリック検定を使用したためであった可能性も十分ある。あくまで今回の結果から判断するのであれば、ベースラインの時点では傾眠率や拒否率、むせ率は認知症と関連が無いようである。ここでなぜか？考えることが重要である。例えば 50%程度あるむせ率は、平均年齢が 85 歳前後であるので、認知症というよりは高齢であることと関係していたのかもしれない。また、有意差はないものの女性の傾眠率は高いように思える。また、女性のみ食事満足度に差が出たことからなども、認知症に関連する変数には男女差があるように見える。次項からは、この対象者が１年間の期間を経て、どのように変わるか調べていく。あまり変わらないかもしれないし、ベースラインで MMSE が低く認知症であることは、予後に影響を与えるかもしれない。

チャレンジ８【C８】
さらに解析を進める前に、必要な情報を整理しよう

01 以下の文章の空欄および選択肢にチェックを入れなさい。【　・　】は正しい文章に〇をつけなさい。

この問題より、ベースラインと１年後の変化について検討していく。

このような解析方法を【横断的解析・縦断的解析】という。また、このような解析にもちいるデータを【対応のないデータ・対応のあるデータ】という。

ベースラインと１年後の平均値または中央値の変化の検定手法で当てはまるものすべてにチェックを入れなさい。

- ☐ スチューデントのｔ検定
- ☐ マンホイットニーU 検定
- ☐ 対応のあるｔ検定
- ☐ 符号付順位和検定

同様に割合や該当率の変化の検定手法で当てはまるものすべてにチェックを入れなさい。

- ☐ カイ二乗検定
- ☐ フィッシャーの正確検定
- ☐ マクネマー検定

02 以下に示す通り、認知症の有無および性別により４グループに分けて、１年間の変化を検討していきたい。それぞれのｎ数およびそのグループのみで解析するための条件式を書きなさい。

認知症	性別	ｎ数	条件式
なし	男性	15	認知症==0 & 性別==0
なし	女性	29	認知症==0 & 性別==1
有病	男性	20	認知症==1 & 性別==0
有病	女性	36	認知症==1 & 性別==1

ポイント

各解析グループの人数を見て、パラメトリック検定を使うか、ノンパラメトリック検定を使うか判断するとよい。

チャレンジ９【C9】
認知症ではない男性は、１年間でどのように変化するか調べてみよう

01 認知症非有病群の男性におけるベースラインと１年後の変化を検定し、以下の表を埋めなさい。

認知症-非有病群【男性】	ベースライン n = 15	１年後 n = 15	*P*
	mean（SD）	mean（SD）	
年齢，歳	86.9(8.4)	87.9(8.4)	<0.001
身長，cm	165.5(6.6)	165.4(6.4)	0.343
体重，kg	62.0(11.3)	62.5(11.6)	0.232
エネルギー摂取量，kcal	1388.3(71.0)	1388.1(70.0)	0.798
BMI，kg/m^2	22.6(3.5)	22.8(3.7)	0.188
	中央値（四分位範囲）	中央値（四分位範囲）	*P*
幸福度，点	3.00 [3.00, 4.00]	3.00 [3.00, 4.00]	0.964
食事満足度，点	4.00 [4.00, 5.00]	5.00 [4.00, 5.00]	0.666
	%	%	*P*
食事中の傾眠率	33.3	33.3	1.000
食事の拒否率	6.7	13.3	1.000
食事中のむせ率	53.3	53.3	1.000

02 １年間で有意に増加または減少した変数があれば書き出しなさい。

有意に増加	有意に減少
年齢	なし

ポイント

年齢～BMI についても、人数が 15 人しかいないので、対応のある t 検定ではなく、符号付順位和検定を使う方が良い。

チャレンジ 10【C10】
認知症ではない女性は、１年間でどのように変化するか調べてみよう

01 認知症非有病群の女性におけるベースラインと１年後の変化を検定し、以下の表を埋めなさい。

認知症-非有病群【女性】	ベースライン n = 29	１年後 n = 29	*P*
	mean（SD）	mean（SD）	
年齢，歳	84.1(6.2)	85.1(6.2)	<0.001
身長，cm	154.6(5)	154.5(5.1)	0.581
体重，kg	49(5.8)	49.2(6.5)	0.560
エネルギー摂取量，kcal	1333.7(64.8)	1330.7(67.0)	0.615
BMI，kg/m^2	20.5(2.4)	20.6(2.6)	0.430
	中央値（四分位範囲）	中央値（四分位範囲）	*P*
幸福度，点	3.00 [3.00, 4.00]	3.00 [3.00, 4.00]	0.268
食事満足度，点	5.00 [4.00, 5.00]	4.00 [4.00, 5.00]	0.186
	%	%	*P*
食事中の傾眠率	6.9	13.8	0.617
食事の拒否率	6.9	13.8	0.480
食事中のむせ率	44.8	51.7	0.617

02 １年間で有意に増加または減少した変数があれば書き出しなさい。

有意に増加	有意に減少
年齢	なし

ポイント

年齢～BMI についても、人数が 29 人しかいない。対応のある t 検定と判断に悩むところだが、 n 数が 30 未満であるため今回の結果は符号付順位和検定を用いた。

チャレンジ 11【C11】
認知症であることが原因で、男性は１年間でどのように変化するか調べてみよう

01 認知症有病群の男性におけるベースラインと１年後の変化を検定し、以下の表を埋めなさい。

認知症-有病群 【男性】	ベースライン n = 20	１年後 n = 20	*P*
	mean（SD）	mean（SD）	
年齢，歳	84.4(8.0)	85.4(8.0)	<0.001
身長，cm	165.7(6.2)	165.6(6.3)	0.586
体重，kg	62.3(10.7)	62.3(11.5)	0.920
エネルギー摂取量，kcal	1417.4(97.5)	1408.8(98.3)	0.030
BMI，kg/m^2	22.6(3.3)	22.6(3.4)	0.729
	中央値（四分位範囲）	中央値（四分位範囲）	*P*
幸福度，点	3.00 [2.75, 4.00]	4.00 [3.00, 4.00]	0.235
食事満足度，点	4.00 [3.00, 4.00]	4.00 [3.00, 5.00]	0.349
	%	%	*P*
食事中の傾眠率	35.0	75.0	0.027
食事の拒否率	10.0	55.0	0.008
食事中のむせ率	55.0	50.0	1.000

02 １年間で有意に増加または減少した変数があれば書き出しなさい。

有意に増加	有意に減少
年齢，傾眠率，拒否率	エネルギー摂取量

チャレンジ 12【C12】
認知症であることが原因で、女性は１年間でどのように変化するか調べてみよう

01 認知症有病群の女性におけるベースラインと１年後の変化を検定し、以下の表を埋めなさい。

認知症-有病群 【女性】	ベースライン n = 36	１年後 n = 36	*P*
	mean（SD）	mean（SD）	
年齢，歳	85.4(7.6)	86.4(7.6)	<0.001
身長，cm	152.6(5.8)	152.8(5.8)	0.172
体重，kg	49.4(8.7)	48.2(9.4)	<0.001
エネルギー摂取量，kcal	1333.6(81.8)	1328.8(81.0)	0.011
BMI，kg/m^2	21.1(3.0)	20.5(3.3)	<0.001
	中央値（四分位範囲）	中央値（四分位範囲）	*P*
幸福度，点	3.00 [2.00, 4.00]	3.50 [2.00, 4.00]	0.248
食事満足度，点	4.00 [3.00, 4.00]	3.00 [2.00, 4.00]	0.001
	%	%	*P*
食事中の傾眠率	25.0	55.6	0.022
食事の拒否率	11.1	38.9	0.004
食事中のむせ率	50.0	52.8	1.000

02 １年間で有意に増加または減少した変数があれば書き出しなさい。

有意に増加	有意に減少
年齢，傾眠率，拒否率	体重，エネルギー摂取量 BMI，食事満足度

チャレンジ 13【C13】
ここまでの結果をもう一度まとめてみよう

01 以下の表の空欄を埋めなさい。

横断的解析【C7】による認知症有病群の様子

【認知症有病群】	男性	女性
有意に高い変数	なし	なし
有意に低い変数	MMSE	MMSE，食事満足度

縦断的解析【９～12】による１年間の変化

【男性の変化】	有意に増加	有意に減少
共通の変化	年齢	なし
非有病群のみの変化	なし	なし
有病群のみの変化	傾眠率，拒否率	エネルギー摂取量

【女性の変化】	有意に増加	有意に減少
共通の変化	年齢	なし
非有病群のみの変化	なし	なし
有病群のみの変化	傾眠率，拒否率	体重，エネルギー摂取量 BMI，食事満足度

チャレンジ 14【C14】
結果をもとに、認知症と関連する症候との関係を考察しよう

01	以下の文章の空欄および選択肢にチェックを入れなさい。【　・　】は正しい文章に〇をつけなさい。

ベースラインにおいて、認知症有病群と非有病群を比較したところ、女性の食事満足度を除いて、BMI 等の体格や食行動、幸福度などすべての変数で【有意差が見られた・有意差が見られなかった】。

つまり、縦断的解析による１年間の変化は、女性の食事満足度を除いて、

□ ベースラインにおいて認知症を有病していたかどうかが原因であると考えられる。

□ ベースラインにおいて有意差が見られなかった変数が原因であると考えられる。

縦断的解析の結果を見ると、認知症の非有病群では男女ともに、１年間で有意な変化が見られた変数はなかった。このことから、

□ 認知症を有病していなくても、１年の加齢により、様々な測定値が悪化する。

□ 認知症を有病していない場合、１年の加齢による変化はない。

認知症の有病群を見ると、１年間で男女ともに（　傾眠率　拒否率　エネルギー摂取量　）が有意に悪化した。女性では加えて（　体重　BMI　食事満足度　）が悪化している。

以上より、今回の研究の結論として、

□ 認知症と上記の悪化した要因との相関関係を示すことができた。

□ 認知症が原因で、１年後に上記の要因が悪化する因果関係を示すことができた。

□ 認知症を有することの１年間の変化を明らかにすることができなかった。

補足とおわりに

ベースラインにおいて食事満足度が女性の認知症有病群で低かったのは、上記の１年間で悪化した因子の中で、先行して低下してきていたからかもしれない。また、女性の方が男性より予後が悪いのも、女性の方が認知症の影響を早く受けるからかもしれない。もしかすると、今後男性においても体重や BMI が低下するかもしれない。この「かもしれない」については、解析の切り口を変える、先行研究の結果を参考にする、新しい研究を計画することで解決していくことになる。

解析の切り口を変える方法の１つに、さらなる統計手法を学ぶ手段がある。引き続き、高度統計を学び、解析の幅を広げていこう。