

レポート課題

解答例

以下に、レポートの解答例を示します。ただし、別解もあり得ます。

※ 間違いを発見された場合は、指摘してくださると助かります。

1. 回帰分析の結果の解釈に関する問題

- (1) “外向性”と“社会的スキル”の得点がともに0点の小学生における“対人ストレス”の予測値が10.5点である。 (5点)

[補足] 切片は、全ての独立変数が0の場合の従属変数の予測値に相当します。

- (2) “社会的スキル”得点が一定であった場合、“外向性”得点が1点高くなれば“対人ストレス”の予測値は1.2点低くなる。 (5点)

[補足] 重回帰分析においては、独立変数同士の相関関係を考慮して偏回帰係数の値を解釈する必要があります。係数の値がマイナスであることに注意しましょう。

- (3) 結論づけられない。“外向性”と“社会的スキル”の単位が異なるため（標準化が行われていないため）、偏回帰係数の値からはいずれがより“対人ストレス”の個人差を左右しているかはわからない。 (5点)

[補足] 標準偏回帰係数は（従属変数と）独立変数がすべて標準化されているので、0点は平均、1点はSD1つ分だという意味において単位がそろいます。ただし、独立変数同士に相関があることが一般的なので、“外向性”と“社会的スキル”以外の独立変数を追加すれば、標準偏回帰係数の値自体が変化する可能性もあります。従って、どの独立変数がより従属変数の個人差を説明しているかは、標準化された結果であったとしても、投入されている独立変数をセットとして、その中での話として考える必要があります。

- (4) 結論づけられない。-3.5という偏回帰係数の値は（“外向性”得点が一定の下での）“社会的スキル”と“対人ストレス”との間の相関関係について述べているだけである。 (5点)

[補足] 一時点での調査データから得られた偏回帰係数は、個人の中での変化を示すものではなく、あくまで個人差を記述したものであることを常に意識することが必要です。

- (5) “外向性”と“社会的スキル”がいずれも集団の平均である小学生における“対人ストレス”の予測値を意味する。 (5点)

[補足] センタリングによって、切片は（外向性-外向性の平均）と（社会的スキル-社会的スキルの平均）がともに0点である小学生における対人ストレス得点の予測値になります。これは、外向性がその平均、社会的スキルがその平均である小学生における対人ストレス得点の予測値に他なりません。テストやアンケートの得点には明確な原点が存在せず、そもそも0点を取り得ないようなものも存在します。（“いいえ”=1点、“はい”=4点の4件法で集計を行う場合など。）こうした場合でも切片に積極的な意味を持たせることができるように、センタリングは有効な手段です。

2. 相関関係の評価に関する問題

- (1) 相関係数では直線的な相関関係しか評価することができないのに対し、記事によればテレビの視聴時間と学力テストの得点の間には曲線的な相関関係があるため。(5点)
- (2) 調査結果が示しているのは相関関係であり、個人の変化を保証するものではないということが平易な言葉で説明されていればよい。(10点)

なお、上記も含めて説明に使える情報としては、以下のような例が挙げられる。

- この調査結果が示しているのは因果関係ではなくあくまで相関関係である。
 - 他の要因が統制されていないので、相関関係すら直接的なものかどうかは判断できない。つまり、まったくテレビを見ない子供と1時間未満見る子供とでは、テレビの視聴時間以外に学力を左右する何かが異なっている可能性がある。(例えば、親子のコミュニケーションの量、家庭の社会経済的地位、ゲームやインターネットなどテレビ以外の娯楽の有無など。)
 - 仮に因果関係が成り立っていたとしても、その関係の強さを評価する情報がない。つまり、1時間未満テレビを見ているのに、まったく見ていない子供よりテスト成績の悪い子供もたくさんいる可能性がある。関係の強さを評価する情報がない限り、個々の事例をどの程度正確に予測できるかはわからない。
 - この調査に用いられたテストそのものが学力を測定するツールとして妥当なものであるという保証がない。
 - テレビの視聴時間は自己報告によるものであり、相関関係すら信頼できる結果であるとは限らない。
- (3) 直線的でない相関関係を示すような事例が、もっともらしい理由とともに挙げられていればOK。(秀逸なものについては、次年度以降の授業で使わせて頂くかもしれません。)(10点)

3. パス解析に関する問題

- (1) “予習時間”、“復習時間”、“知能指数”(5点)

[補足] 単方向のパスが一度も自分の方に向いていない変数が外生変数です。

- (2) “予習時間”から“授業理解度”へのパス係数は、“復習時間”が一定であった場合、“予習時間”が1SD多くなると、“授業理解度”の予測値は0.2SD高くなる、ということを意味する。“知能指数”から“授業理解度”へのパス係数は、“知能指数”が1SD高くなると、“授業理解度”の予測値は0.2SD高くなる、ということを意味する。このような解釈の違いが生じるのは、“予習時間”が“復習時間”との間に $r=.3$ の相関関係があるのに対し、“知能指数”は他の独立変数と無相関であることが仮定されているためである。(10点)

[補足] 重回帰分析の場合と同様、パス係数の値を解釈する際にも独立変数同士の相関を考慮に入れる必要があります。特に、パス解析では独立変数同士に想定されている相関関係が重回帰分析に比べて複雑である(というか見えにくい)場合が多いので、特に意識する必要があります。また、この例は標準化解であるので、パス係数の値はすべて標準偏差(SD)を単位として行われることとなります。標準化解において何点高い(多い)というのは、「標準偏差にして何個分高い(多い)」という意味です。

(3) $0.20 \times 0.25 + 0.10 = 0.15$; “知能指数”が1SD高くなると、“テスト成績”の予測値は0.15SD高くなる。(10点)

[補足] この場合の総合効果は、“知能指数”と“テスト成績”間の単回帰的な関係となります。

4. 因子分析の実行と結果の解釈に関する問題 (25点)

Rを用いて因子分析を実行するスクリプト例は以下の通り。ただし、データファイルはデスクトップに保存し、psychパッケージ内の関数fa()によって最小2乗法、プロマックス回転による因子分析を行ったものとする。因子数は、相関行列の固有値とその減衰状況から2が最適であると判断した。

```
# データの読み込み
data01 <- read.csv("~/Dropbox/Users/User/kimatsu.csv")

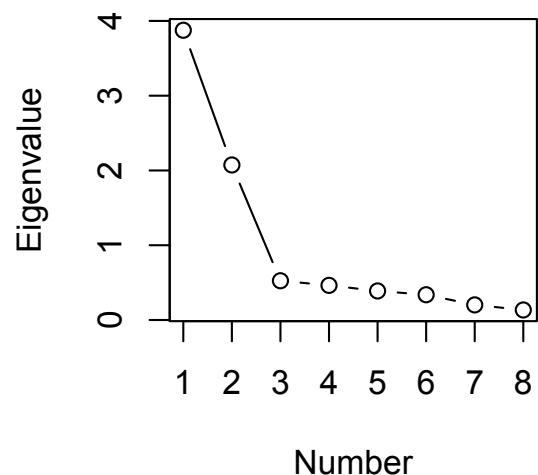
# 相関行列の固有値の算出、スクリープロットの描画
eigen01 <- eigen(cor(data01))$values; eigen01
plot(eigen01,type="b",xlab="Number",ylab="Eigenvalue")

# psychパッケージの読み込みとfa()による因子分析の実行
library(psych)
fit01 <- fa(r=data01, nfactors=2, rotate="promax", fm="minres")
print(fit01)
```

相関行列の固有値は以下の通り。それをプロットしたものが右下の図である。

```
[1] 3.8754105 2.0735495 0.5250003 0.4633091 0.3885293 0.3378484 0.2021201 0.1342327
```

項目	因子1	因子2	共通性
明るい	0.84	0.01	0.71
生き生きしている	0.63	0.10	0.45
楽しい	0.88	-0.16	0.70
元気いっぱいだ	0.92	0.08	0.90
仲がよい	0.16	0.76	0.68
まとまりがある	-0.12	0.85	0.67
きっちりしている	0.01	0.77	0.60
思いやりがある	-0.03	0.77	0.58
寄与	2.75	2.53	
寄与率	0.34	0.32	
因子間相関		0.34	



因子分析の結果は上の表の通り。第1因子に強い負荷を持つ項目の内容、第2因子に強い負荷を持つ項目の内容から、例えば第1因子を“活発性”、第2因子を“凝集性”などと命名できるだろう。

5. (略) (25点)