

# 心理学論文における数値と統計の書き方

2018年7月9日

新美亮輔(新潟大学人文学部)

## 目次

### 1 桁数を決めて書く

- 1.1 数値は3桁か4桁にそろえる
- 1.2 「小数点以下〇桁」も意識する
- 1.3 適切な単位を選ぶ
- 1.4 略してよいゼロと略してはいけないゼロ
  - 末尾のゼロ
  - 小数点の前のゼロ
- 1.5 計算途中では四捨五入しない
- 1.6 Excel や統計ソフトで数値の最後につく“E+5”とか“e-10”の意味

### 2 統計的検定結果の書式

- 2.1 実例を見るのがいちばん
- 2.2 本文中での統計数値の典型的な書き方
  - 平均値など
  - z 検定
  - t 検定
  - 分散分析
  - $\chi^2$  検定
  - 相関係数
- 2.3 書式の要注意ポイント
  - 何でもかんでもイタリック体にしない
  - Century は非推奨
  - 半角スペースをたくさん入れる
  - カイ2乗のカイをエックスで代用しない
  - 検定統計量は臨界値（境界値）ではなくデータから実際に算出された値を書く
- 2.4  $p$  値の書き方
  - $p > .05$  とか  $p < .01$  ではなく実際の  $p$  値を書く
  - 大文字 ( $P$ ) か小文字 ( $p$ ) か
  - 不等号の向きに注意
  - 有意でないことを表す  $ns$  は本文中では特に必要ない

有意水準を表すときは  $p$  ではなく  $\alpha$   
ときどき見かける “ $ps$ ” の意味

### 2.5 Excel で $p$ 値を求める

$z$  検定 (両側検定)

$t$  検定 (両側検定)

分散分析 ( $F$  値)

$\chi^2$  検定

### 2.6 分散分析表は必要?

## 3 統計的検定結果の説明のしかた

### 3.1 複数の検定結果を列挙するとき

カッコに入れ, セミコロン ; で区切るのが基本

悪い例 (1) 文中での列挙

悪い例 (2) 箇条書き

悪い例 (3) テンプレ文の繰り返し

### 3.2 検定結果はあくまでも言葉で説明する

### 3.3 「考察」で何を書けばいいか

どこにも有意差がなかった場合

予想外の有意差があった場合

論文の本題とはあまり関係がないところに差があった場合

※ この資料は心理学専攻の大学生向けです。他の分野にはあてはまらないことが多数ありますので、ご注意ください。

※ この資料は心理学論文の「正しい」書き方を定義するものではありません。大学生が心理学論文の「適切な」書式を学ぶための資料です。ジャーナルによって、また研究者によって、使う書式が異なることもあります。授業レポートでは、担当教員の指示に従って下さい。

## 1 桁数を決めて書く

### 1.1 数値は3桁か4桁にそろえる

数値は本文でも表でも適切な桁数にそろえて書きます。心理学論文では、たいてい3桁か4桁（ときどき2桁）です。つまり、3桁か4桁になるように四捨五入します（四捨五入することを英語で round というので、「丸める」とも言います）。ただし、計算途中では四捨五入しません（1.5 参照）。最終的に論文に書く時にだけ四捨五入します。

- × 123.71064981 → ○ 124 または 123.7
- × -4.198350112 → ○ -4.20 または -4.198

同じ種類の数値は同じ桁数で書くのが原則です。違う種類の数値ならば桁数は違っていてもかまいません。たとえば正答率 (%) は3桁、反応時間 (ms) は4桁、というのはOKです。

- × 正答率は自転車条件で 68%，飛行機条件で 82.9% だった。
- 正答率は自転車条件で 68.2%，飛行機条件で 82.9% だった。
- 正答率は自転車条件で 68%，飛行機条件で 83% だった。
  
- 新幹線条件の平均正答率は 68.2%，平均反応時間は 1234 ms だった。

データによっては、同じ種類の数値でも3桁の数（921）と4桁の数（1127）が混在することもあります。これはかまいません。無理に揃えることはしません。

- 平均反応時間は通勤電車条件で 862 ms，特急電車条件で 1059 ms だった。
- × 平均反応時間は通勤電車条件で 862.4 ms，特急電車条件で 1059 ms だった。

なぜ 862.4 と 1059 がダメなのかと言うと、桁数はデータの精度も考慮して決めるものだからです。通常の PC とキーボードを使った反応時間測定の場合、1 ms を下回るような細かい時間はさほど正確に測定できない（誤差の範囲内）と考えられます。それに、たいていの心理学実験では、人間の反応時間の 1 ms 程度の差は小さすぎてあまり意味がありません。ですから、ms 単位の反応時間ならば1の位まで書けば十分なわけです。

### 1.2 「小数点以下○桁」も意識する

このように、実際のデータの表記は、小数点の位置も考慮した桁数で書きます。測定精度や計算精度、実質的に意味があるかを考慮し、桁数が3、4桁程度におさまるよう、小数点以下何桁まで書くのかを決めましょう。

分散分析の  $F$  値や  $t$  検定の  $t$  値などの検定統計量は、通常の心理学実験のデータでは小数点以上が1~2桁になるのが普通ですから、小数点以下1、2桁程度まで書きます。 $p$  値は1の位には意味がない（略してしまうことも多い）ので、小数点以下3桁まで書くことが多いですね。相関係数は小数点以下2桁まで書くことが多いようです。『APA マニュアル』は、

相関係数、割合、 $F$ 、 $t$ 、 $\chi^2$ などは小数点以下2桁、 $p$ 値は小数点以下2または3桁、と指定しています。なお、 $r = .40$ のような末尾のゼロを略してはいけません（1.4参照）。

### 1.3 適切な単位を選ぶ

0.000000538 km/h とか 12501506904 秒 のような数値は桁が多すぎて不便です。桁数が3, 4桁になるよう、単位を変えましょう。日本心理学会の『執筆・投稿の手びき』（以下『手びき』）は、数値が0.1から1000の範囲になるように単位を選ぶとしています。

× 0.000288 km → ○ 28.8 cm または 0.288 m

× 445091 Hz → ○ 445.1 kHz（キロヘルツ）または 445 kHz

心理学実験の刺激・装置の大きさ・長さは、たいてい1 mm（0.001 m）単位で測ると思います（メジャーには1 mm単位で目盛がついていますよね）。その場合、28.8 cm とか 0.288 m は測定の精度という観点からも妥当な表記です。一方、1 mm単位でしか測定していないのに4.568 cm と書くのはやや過剰です。この場合は、4.6 cm と2桁の表記にしても問題ありませんし、少し詳しくに45.7 mm でもいいでしょう。いずれにせよ、論文の中では同じ種類の数値はできるだけ単位を統一しましょう。刺激の大きさのある場所ではcm単位で、別の場所ではmm単位で書いていたりすると、とても理解しづらくなります。

金額や人口など、どうしても大きい桁数で書く必要がある場合は、桁区切りのコンマを入れて12,345,678 と書くと見やすくなります。桁区切りは3桁ごとに入れます。ただし、「億円」とか「万人」を単位にして桁数を減らすこともできます。自分の論文の目的に照らして適切で読みやすい表記を選んで下さい。

### 1.4 略してよいゼロと略してはいけないゼロ

#### 末尾のゼロ

論文では、安易に略してはいけません。学校の算数や数学では12と12.0は同じですが、実験・調査論文では意味が違います。なぜならば、

12 cm → 実際の測定値は11.500... cm から 12.499... cm の間

12.0 cm → 実際の測定値は11.9500... cm から 12.0499... cm の間

という意味だからです。精度が違うのです。もし1 mm単位で測定したならば、12.0 cm と書くことには意味があります。同様に、たとえば正答率（正答数÷全試行数）ならば、

0.67 → 実際に算出された正答率の値は0.66500...から0.67499...の間

0.670 → 実際に算出された正答率の値は0.669500...から0.670499...の間

という意味になります。現実のデータを扱うときには、誤差や精度の問題が必ずつきまとうということを覚えておきましょう。

Excel が勝手にゼロを略してしまう場合には、右クリックして「セルの書式設定」→「表示形式」タブで「数値」を選び、「小数点以下の桁数」を指定して下さい。

### 小数点の前のゼロ

0.12 のような小数点の前のゼロは、比率（割合）・確率・相関係数・信頼性係数など定義上 1 を超えることがありえない数値に限って略することができます。1 を超える可能性がある数値では略してはいけません。

$$\begin{aligned} & \bigcirc p = .345 \quad \bigcirc r = -.67 \\ & \times t(50) = .89 \quad \rightarrow \quad \bigcirc t(50) = 0.89 \quad (t \text{ 値は } 1 \text{ を超える可能性がある}) \end{aligned}$$

小数点の前のゼロは略さなくても OK ですが、心理学論文では略することが多いようです。いずれにせよ、略すなら略す、略さないなら略さないで統一すること。ただ、図表中の数値に限っては小数点前のゼロを略すが文章中では略さない、という書き方もあります。これは、文章中だと .23 より 0.23 の方が見やすいという考えからだと思われまます。

### 1.5 計算途中では四捨五入しない

四捨五入して桁数をそろえるのは、最終的に論文に書くときだけです。Excel で集計しているときなど、計算の途中ではいちいち四捨五入しません。何度も四捨五入をすると、誤差（いわゆる「丸めの誤差」）が積み重なって最終的な計算結果がかなり変わってしまうことがあるからです。たとえば、 $5 \div 3 \times 12.5$  という計算を考えます。

#### × 途中で四捨五入する場合

$$\begin{aligned} 5 \div 3 &= 1.666666... \rightarrow 1.67 \quad (\text{ここで } 1 \text{ 回四捨五入}) \\ 1.67 \times 12.5 &= 20.875 \rightarrow \mathbf{20.88} \end{aligned}$$

#### ○ 途中で四捨五入しない場合

$$\begin{aligned} 5 \div 3 &= 1.666666... \quad (\text{四捨五入しない}) \\ 1.666666... \times 12.5 &= 20.833333... \rightarrow \mathbf{20.83} \end{aligned}$$

0.05 の差が出ています。実際の集計や分析ではもっと何回も計算が行われますから、誤差はさらに大きくなる可能性があります。

厳密に言えば、Excel などのソフトが行っている計算も、こういう誤差を含んでいます。が、コンピュータは大きい有効桁数で数値を扱っている（Excel ではだいたい 15 桁）上に、誤差を減らすいろいろな工夫がされているので、通常の用途では誤差が小さすぎて気にならないだけです。現実のデータの計算には、多かれ少なかれ、常に誤差がつきまとうのです。

### 1.6 Excel や統計ソフトで数値の最後につく“E+5”とか“e-10”の意味

ソフトの出力結果で、ときどき数値の後に E+5 とか +e8 とか ついていることがあります。これは、桁数がとても多い数値を表すための方法です。意味は下のとおり。

$$\begin{aligned} 1.203\text{E}+5 &= 1.203 \times 10^5 = 1.203 \times 100000 = 120300 \\ 1.203\text{E}-5 &= 1.203 \times 10^{-5} = 1.203 \times 0.00001 = 0.00001203 \end{aligned}$$

つまり指数で表しているわけです（E は指数 exponential の意味）。ソフトからのコピペで論

文に 1.203E+5 と書かないように。こんな桁数の大きい値は、単位を変えるなどして桁数の小さい表記に改めます。統計ソフトでは非常に小さい  $p$  値がよくこの指数表記で出てきますが、論文では 0.001 より小さい  $p$  値は単に  $p < .001$  と書けば OK です (2.4 参照)。人口や金額など、どうしても大きい桁数で書く必要がある場合は、 $9.807 \times 10^8$  のように「 $\times 10^n$ 」式の指数表記にするか、980,700,000 のように桁区切りのコンマ (3 桁ごと) を使うと読みやすくなります。

こういう非常に桁数の多い数値は、まれにですが、コンピュータがうまく扱いきれず、計算ができなくなってしまうことがあります。もしデータの集計や分析の途中でこういう値が出てきていて、ソフトがエラーを出していたら注意しましょう。単位を変える、かけ算や割り算の順序を変える、などの工夫で回避できます。

## 2 統計的検定結果の書式

### 2.1 実例を見るのがいちばん

統計的検定はとて多くの種類があるので (*100 statistical tests* という本があるくらいです)、あらゆる検定法の結果の書き方を規定したマニュアルは存在しません。ですから、書き方に困ったら『手びき』や『APA マニュアル』だけでなく、実際の論文でどう書かれているかを見てみましょう。卒論なら、似た実験・分析をしている先行研究の論文をいくつか読んでいしょうから、そこで検定結果がどう書かれているかを参考にすればよいわけです。

とりあえず、基本形はこんな感じですね。検定結果はまず文で説明し、カッコ内に検定統計量、自由度、 $p$  値、効果量などを書きます。

乗り物 (鉄道・飛行機・船) と移動時間 (短・長) によって旅行の主観的満足度が異なるかを、参加者間 2 要因分散分析によって検討したところ、乗り物 ( $F(2, 34) = 18.94, p < .001, \eta_p^2 = .53$ )、移動時間 ( $F(1, 34) = 5.71, p = .023, \eta_p^2 = .14$ ) とともに主効果が有意だった。さらに、交互作用も有意であった ( $F(2, 34) = 3.30, p = .049, \eta_p^2 = .16$ )。

あくまでも文が基本だということに注意して下さい。統計は添え物としてカッコ内に書くのが基本です。検定統計量や  $p$  値をただ羅列したり、箇条書きにしたりしてはいけません。どうしても煩雑な場合は、図表にしましょう。たとえば、相関行列のすべての相関係数の有意性検定の結果を文中に羅列することはせず、表にまとめるのが一般的です。

### 2.2 本文中での統計数値の典型的な書き方

本文でカッコ ( ) の中に統計的数値類を書くときの典型的な例を挙げます。あくまでも典型的な例なので、こう書かなければいけないわけではありません。また、図表の中での書き方は別ですので、ご注意。

## 平均値など

$M = 20.9, SD = 1.55$	(平均と標準偏差)
$M = 20.9, SD = 1.55, N = 43$	(サンプルサイズを加えた場合)
$M = 20.9 \text{ cm}, SD = 1.55$	(必要なら単位をつける)
$M = 12.3, 95\% \text{ CI} = [11.9, 12.7]$	(平均と 95%信頼区間)
$Mdn = 0.49$	(中央値)

サンプルサイズは、データ全体の総サンプルサイズを大文字の  $N$  で、データの一部のサブサンプルサイズ（たとえば、参加者をいくつかの群に分けたときの 1 つの群の参加者数）を小文字の  $n$  で表すことがあります。

## z 検定

$$z = 2.10, p = .036$$

自由度はありません。両側検定では、 $z$  値はマイナスを略し絶対値を書くことが多いです。

## t 検定

$$t(12) = 3.68, p = .003$$
$$t(12) = 3.68, p = .003, d = 1.06 \quad (\text{効果量として Cohen の } d \text{ を加えた場合})$$

カッコの中に自由度を 1 つ書きます。自由度をカッコから出して

$$t = 3.68, df = 12, p = .003$$

と書くこともあります。心理学ではあまり一般的ではありません。ウェルチの検定（等分散を仮定しない 2 標本の  $t$  検定）では、例外的に自由度を小数で書くことがあります。

$z$  検定と同じく、両側検定では  $t$  値のマイナスを略し絶対値を書くことが多いです。

$t$  検定にはたくさんの種類があるので、どんな  $t$  検定なのかを必ず文中で書きましょう。特に、対応のある検定をしたことが書かれていないレポートをよく見かけます。特に書かれていなければ、多くの読者は対応のない検定だと解釈してしまいます。

## 分散分析

$$F(3, 69) = 1.53, p = .21$$

自由度は 2 つあります。自由度の順番には意味があるので、逆にしないように。1 つ目が  $F$  の分子の自由度、2 つ目が分母（誤差項）の自由度です。

$$F(3, 69) = 1.53, p = .21, \eta_p^2 = .06 \quad (\text{効果量として偏イータ 2 乗を加えた場合})$$

分散分析では、効果量として  $\eta^2$ （イータじじょう）、 $\eta_p^2$ （偏イータじじょう）、 $\eta_G^2$ （一般化イータじじょう）などが使われます。どれでなければいけない、と決まっているわけではありませんが、同じ論文の中ではどれかに統一した方がよいでしょう。

$$F(3, 69) = 1.53, p = .21, MSe = 8.78 \quad (\text{MSe を加えた場合})$$

分散分析では  $MSe$ （分散分析表の誤差項の  $MS$  のこと、 $MSE$  と書くこともある）を加

えることがあります。効果量と *MSe* の両方を書いてもかまいません。

なお、多重比較の結果を書くときは、多重比較の手法の名前（テューキーの HSD 法、シェフェ法、ライアン法、ボンフェローニ法……）も忘れずに明記しましょう。

### $\chi^2$ 検定

$$\chi^2(3) = 9.01, p = .029 \quad (\text{自由度をカッコ内を書く})$$

$$\chi^2(3, N = 62) = 9.01, p = .029 \quad (\text{サンプルサイズも書く場合})$$

### 相関係数

$$r = -.23, p = .037$$

$$r = -.23, N = 80, p = .037 \quad (\text{サンプルサイズも書く場合})$$

有意な相関かどうかという検定（無相関検定）の *p* 値を加えることが多いです。相関係数がいくつかある場合には、区別するために、*r*<sub>HL</sub>（High 条件と Low 条件の間の相関）、*r*<sub>WM</sub>（ある従属変数と作業記憶 *working memory* 課題の成績との相関）のように記号を作ってもかまいませんが、必ず初出時に記号の意味を定義しましょう。

なお、「相関係数」と名前をつくものは、ここで例を挙げている *r*（ピアソンの積率相関係数、もっとも代表的な「相関係数」）以外にもありますので、注意しましょう。

## 2.3 書式の要注意ポイント

### 何でもかんでもイタリック体にしない

数値、カッコ、等号や不等号はイタリック体にしません。統計的概念の記号はイタリック体にします。平均の *M* や標準偏差の *SD* をはじめとする種々の統計量 (*z*, *t*, *F*, *MSe*, *df*, *r*, *R*<sup>2</sup>, *d'* などなど)、サンプルサイズの *N* や *n*、確率の *p* などです。数式で変数名として使う *a*, *b* などもイタリック体にする人が多いです。ギリシャ文字はイタリック体にしなくても大丈夫です（検定統計量の  $\chi^2$ 、有意水準や信頼性係数の  $\alpha$ 、標準偏回帰係数の  $\beta$ 、効果量の  $\eta^2$  など）。

イタリック体にしない統計記号・略語も多数あります。何かの数値を表すわけではない分析手法名や関数名などは立体（イタリック体でない）にします。たとえば ANOVA（分散分析）、GLM（一般化線形モデル）、log（対数）など。『APA マニュアル』では、統計量である平均の標準誤差（standard error of mean）は *SEM* とイタリック体、分析手法である構造方程式モデリング（structural equation modeling）は *SEM* と立体にしています。テューキーの HSD も、検定手法の名前として書く場合は立体でよいでしょう。

このように、何らかの数値を表しうる統計量・スコア・変数等の名前はイタリック体にすることが多いですが、例外もあります。たとえば、『APA マニュアル』では CI（信頼区間）、AIC（赤池情報量規準）などは立体です。迷ったら論文の実例や『APA マニュアル』表 4.5 を参考にしましょう。基本的には、イタリック体にする確信があればイタリック体にし、迷ったら立体にしておく、という方針でよいと思います。

### Century は非推奨

Word のデフォルトの英数字フォントである Century はイタリック体にならず、ただの斜体にしかならないので、論文には不向きです。イタリック体は、単に文字を斜めにした斜体



(ax) ではなく、筆記体に近い字形 (ax) になります。イタリック体が使えるフォントとして、Times New Roman がよく使われます。

△	$r = .70, p < .001$	$df = n - 1$	(Century)
△	$r = .70, p < .001$	$df = n - 1$	(MS 明朝)
○	$r = .70, p < .001$	$df = n - 1$	(Times New Roman)
○	$r = .70, p < .001$	$df = n - 1$	(Century Schoolbook)

#### 半角スペースをたくさん入れる

半角スペースをまったく入れていないレポートをよく見かけます。スペースがないと、下のように文字が詰まって非常に読みづらくなります。

×  $F(5,60)=2.74,p<.05$     × 移動速度は平均 1.03deg/s( $SD=0.04$ )だった。

半角コンマの後、等号や不等号の前後、数値と半角の単位の間などには半角スペースが必要です。スペースを使うのは英語の基本中の基本ですから、半角で書いている時は必ず半角スペースを使って下さい（これは英語文献の表記にも共通の注意点です）。

○  $F(5,60) = 2.74, p < .05$     ○ 移動速度は平均 1.03 deg/s ( $SD = 0.04$ ) だった。

半角スペースが必要な場所を色で示すと、

$F(5,60) = 2.74, p < .05$     移動速度は平均 1.03 deg/s ( $SD = 0.04$ ) だった。

となります。 $F$ と半角カッコ（の間は、フォントによっては詰まってしまうので、適宜スペースを入れて下さい（または、カッコだけは全角を使うという手もあります）。APA スタイルでは  $\chi^2(5)$  のように検定統計量の記号とカッコの間にあまりスペースがないですが、『心理学研究』では  $\chi^2(5)$  のようにややスペースが空いています。これはどちらでもかまいませんが、ひとつのレポートの中では統一しましょう。

次のケースでは、スペースを入れません。

- ・ 数の桁区切りの半角コンマの後    3,000 ms    (× 3,000 ms)
- ・ 小数点として使う半角ピリオドの後     $SD = 12.3$     (×  $SD = 12.3$ )
- ・ 負の数値を表すマイナス符号の後     $r = -.78$     (×  $r = - .78$ )
- ・ 数値と%または角度の単位 (° ' ") の間    12.3% 45.6°    (× 12.3 % 45.6°)

#### カイ 2 乗のカイをエックスで代用しない

$\chi^2$  を  $x^2$  とか  $x^2$  としてはいけません。これは決して細かいことではなく、「ろ」を「3」で代用するくらいダメな書き方です。必ずギリシャ文字のカイを使うこと。「かい」と打って漢字変換（スペースキー押し）すると全角文字の「χ」が出てきますから、これを選択してフォントを Times New Roman にすると、半角の  $\chi$  になります。また、半角の  $\chi$  は Word の「挿入」→「記号と特殊文字」から直接入力できます。

#### 検定統計量は臨界値（境界値）ではなくデータから実際に算出された値を書く

$t$  値や  $F$  値などの検定統計量は、実際にデータから算出された値を書きます。臨界値（ち

ようど  $p$  値が .05 とか .01 になる値、つまり統計の教科書の後ろの数表に載ってるやつ) は書きません。たとえば、 $t$  検定をして  $t$  値が 3.45 (自由度 20) と算出されたとします。有意水準 5% で両側検定の場合、自由度 20 で  $t$  の臨界値は 2.09 で、実際の  $t$  値 (3.45) はこれを上回っているので有意と判断することになります。この場合の書き方は、

- $t(20) = 3.45, p < .05$  (データから算出された  $t$  値を書く)
- ×  $t(20) = 2.09, p < .05$  (臨界値を書くのは誤り)
- ×  $t(20) = 3.45 > 2.09, p < .05$  (不等号で臨界値を書くことはしない)

です。 $t$  の臨界値は誰でも調べればすぐにわかるので、書く必要はないのです。 $z, F, \chi^2$  など他の統計量を使った検定でも同じです。

## 2.4 $p$ 値の書き方

$p > .05$  とか  $p < .01$  ではなく実際の  $p$  値を書く

$p$  値を不等号で書くのは、古い書き方です。統計ソフトが一般的ではなかった時代には、実際の  $p$  値がわからず、検定表で臨界値を調べて検定するしかなかったためです。現在では Excel や統計ソフトで実際の  $p$  値が簡単にわかる (2.5 参照) ことが多いですから、データから計算された実際の  $p$  値を書くことが推奨されています。なお、不等号で表す時は、有意水準として一般的な値 (.05, .01, .10) を使って書くこと。

- △  $t(20) = 2.22, p < .05$       $F(2, 34) = 1.23, p > .10$      (有意水準と比べる書き方)
- ×  $t(20) = 2.22, p < .04$       $F(2, 34) = 1.23, p > .3$      (有意水準ではない)
- $t(20) = 2.22, p = .038$       $F(2, 34) = 1.23, p = .305$      (具体的な  $p$  値を書く)
- ×  $t(20) = 2.22, p = .038 < .05$       $F(2, 34) = 1.23, p = .305 > .10$      (くどい)

ただし、実際の  $p$  値が .001 より小さい場合は、不等号を使って  $p < .001$  と書きます。

- ×  $t(20) = 5.67, p = .0000015$  (0 が多くて読みづらい)
- ×  $t(20) = 5.67, p = .000$
- $t(20) = 5.67, p < .001$
- ×  $t(20) = 5.67, p < .00001$  (心理学論文で用いられる有意水準は最小で .001)
- △  $t(20) = 5.67, p < 10^{-3}$  (間違いではないが、心理学では一般的でない表記)

なお、 $p$  値は小数点以下 2 桁または 3 桁に四捨五入するのが普通です。また、Excel や統計ソフトで  $p$  値が 5.48E-05 のように E つきの指数表記 (1.6 を参照) になっている場合は、そのまま書いてはいけません (たいていは  $p$  値が非常に小さいケースだと思いますので、 $p < .001$  になると思います)。

大文字 ( $P$ ) か小文字 ( $p$ ) か

心理学分野では小文字の  $p$  の方が一般的です。『手びき』や『APA マニュアル』でも小文字です。大文字を使って ( $t(12) = 3.68, P = .003$ ) のように書くジャーナルもありますが、レポートや卒論では特段の理由がない限り小文字にしておけばよいでしょう。

### 不等号の向きに注意

実際の  $p$  値がわからず、不等号で  $p$  値を報告する場合は、不等号の向きに注意して下さい。逆になっているレポートをけっこう見かけます。 $t$  検定や  $z$  検定の両側検定、分散分析や  $\chi^2$  検定の場合、

$t, z, F, \chi^2$  (の絶対値) が 5% 臨界値より大きいなら、5% 水準で有意なので、 $p < .05$

$t, z, F, \chi^2$  (の絶対値) が 5% 臨界値より小さいなら、5% 水準で有意でないので、 $p > .05$

です。検定統計量 ( $t, z, F, \chi^2$ ) が臨界値より大きいときは、 $p$  値は有意水準より小さいという関係です。

### 有意でないことを表す ns は本文中では特に必要ない

統計的に有意でないことを表す  $ns$  という略語がありますが、論文の本文中ではあまり使うべき理由がありません。 $ns$  だけだと、具体的な  $p$  値も有意水準もわからず、情報量が少ないからです。ただし、ざっくりと簡単に結果を書きたいときや、図表などでは、省スペースのために  $ns$  を使うことがあります。

○  $t(20) = 0.45, p = .66$

△  $t(20) = 0.45, p > .10$  (具体的な  $p$  値を書いた方がよい)

△  $t(20) = 0.45, ns$  (具体的な  $p$  値を書いた方がよい)

△  $t(20) = 0.45, p = .66, ns$  (くどい)

### 有意水準を表すときは $p$ ではなく $\alpha$

5% 水準で検定したことを表すために  $p = .05$  と書いている例を見かけますが、誤りです。 $p$  は実際にデータから得られた  $p$  値を表す記号です。有意水準を明記したい場合は、

○ ……について  $z$  検定 ( $\alpha = .05$ ) を行ったところ、……

○ ……かどうか有意水準 5% で  $z$  検定を行ったところ、……

のように、 $\alpha$  (アルファ) で表すか、言葉で説明します。次のような書き方はしません。

× ……について  $z$  検定 ( $p = .05$ ) を行ったところ、……

× ……かどうか有意水準  $p = .05$  で  $z$  検定を行ったところ、……

× ……に統計的な有意差は見られなかった ( $z = 1.23, p = .05$ )。

両側  $z$  検定で  $z = 1.23$  なら、 $p$  値は  $.22$  になりますから、上の例では ( $z = 1.23, p = .22$ ) と書くのが正解。もちろん、 $\alpha = .05$  で有意差はありません。

### ときどき見かける “ps” の意味

何か特殊な意味のある略語ではなく、 $p$  の複数形 ( $p$  ず) です。複数の検定結果をざっくり簡単に報告するときに使うことがあります。つまり、 $ps > .05$  と書いてあれば、複数の  $p$  値があっただけで、どれも  $.05$  より大きかったよ、という意味です。

たとえば、ある課題の正答率が、参加者の特性を表す 8 つの尺度 (年齢, VIQ, PIQ, 作業記憶容量……) のいずれとも有意な相関を示さなかった、というときに、それが論文の中で

あまり重要な結果ではなく、8個の相関係数とその検定結果をいちいち全部書くほどでなければ、「正答率との相関はいずれの尺度でも非常に小さく、有意ではなかった ( $p > .05$ )。」のように簡単に書いてすませることがあります。

同じように、複数の相関係数を  $r_s$  と表すこともあります。ちなみに、 $p_s$  でも  $r_s$  でも  $s$  はイタリック体にしません ( $p$  と  $r$  はイタリック体)。

## 2.5 Excel で $p$ 値を求める

下のように、たいていの  $p$  値は Excel 関数で計算できます (統計数値表に載っている 5% 臨界値を入力して、約 0.05 になるか試してみてください)。

### $z$ 検定 (両側検定)

$z$  値が 2.3057921 の場合は `= (1-NORM.S.DIST (2.3057921, TRUE)) * 2`

( $z$  値は絶対値をとります。-2.3057921 の場合は、2.3057921 として計算します。絶対値を求める Excel 関数は ABS () です。)

### $t$ 検定 (両側検定)

$t$  値が 1.9143057、自由度 18 の場合は `=T.DIST.2T (1.9143057, 18)`

( $t$  値は絶対値をとります。-1.9143057 の場合は、1.9143057 として計算します。Excel 関数 T.DIST.2T () は  $t$  値に負の値を指定するとエラーになるので注意。)

### 分散分析 ( $F$ 値)

$F$  値が 3.0402915、分子の自由度 2、分母 (誤差項) の自由度 38 の場合は  
`=F.DIST.RT (3.0402915, 2, 38)`

( $F$  値が負になることはありません。2つの自由度の順番には意味があるので、逆にしないように。論文で書く順番 ( $F (2, 38)$ ) と同じです。)

### $\chi^2$ 検定

$\chi^2$  値が 0.8419956、自由度 4 の場合は `=CHISQ.DIST.RT (0.8419956, 4)`

$t$  値などの検定統計量は論文に書くときは四捨五入しますが、上のように  $p$  値を計算するときは四捨五入せず、桁数が多いまま代入します (丸めの誤差を減らすため。1.5 参照)。Excel では、検定統計量が計算してあるセルを直接指定して `=T.DIST.2T (F15, G15)` のように入力するとよいでしょう (セル F15 に  $t$  値が、G15 に自由度が計算してある場合)。

なお、上で紹介している Excel 関数は古いバージョンの Excel では使えないことがあります。古いバージョンの Excel では、NORMSDIST, TDIST など別の名前の (古い) 関数を使います。数値の書き方も違いますから、注意して下さい。

## 2.6 分散分析表は必要?

通常、論文には分散分析表は載せません。ただ、授業のレポートでは統計の勉強のために (きちんと検定ができているか確認できるように) 分散分析表を載せろという指示が出るこ

とがあります。その場合は、レポートの本文中に表として挿入してもいいですし、レポート末尾に付録としてまとめることもあります。

なお、論文に分散分析表を載せるときは、Excel や統計ソフトの出力をそのままコピーしただけのものはダメです。表として適切な書式になるよう、体裁を整えて下さい。

#### × コピーそのままの分散分析表

Table of Analysis of Variance

source	SS	df	MS	F	p
subject	15606.1333333	9	1734.0148148		
A:Factor A	22210.0666667	2	11105.0333333	3.725	0.0443 *
error[AS]	53657.2666667	18	2980.9592593		
Total	91473.4666667	29			

+ p<.10, \* p<.05, \*\* p<.01, \*\*\* p<.005, \*\*\*\* p<.001

#### ○ 適切な書式の分散分析表

表 2. 実験 1 反応時間の結果についての分散分析表

変動因	SS	自由度	MS	F
参加者	15606	9	1734	
乗り物	22210	2	11105	3.73 *
誤差	53657	18	2981	
全体	91474	29		

Note. \*  $p < .05$ .

### 3 統計的検定結果の説明のしかた

#### 3.1 複数の検定結果を列挙するとき

カッコに入れ、セミコロン ; で区切るのが基本

- その結果、乗り物と移動時間のどちらも主効果が有意だった（それぞれ  $F(2, 34) = 18.94, p < .001, \eta_p^2 = .53$ ;  $F(1, 34) = 5.71, p = .023, \eta_p^2 = .14$ ）。
- ……3 要因分散分析の結果、いずれの主効果も有意ではなかった（乗り物,  $F(2, 204) = 1.89, p = .15$ ; 性別,  $F(1, 102) = 1.94, p = .17$ ; 年齢,  $F(2, 102) = 1.04, p = .36$ ）。

### 悪い例 (1) 文中での列挙

- × ……3 要因分散分析を行った。その結果、乗り物の主効果  $F(2, 204) = 5.67, p = .004$ , 性別の主効果  $F(1, 102) = 19.42, p < .001$ , 年齢の主効果  $F(2, 102) = 4.12, p = .019$  となった。

列挙するならば、必ずカッコの中に入れてみましょう。

- ……3 要因分散分析を行った。その結果、いずれの主効果も有意となった（乗り物,  $F(2, 204) = 5.67, p = .004$ ; 性別,  $F(1, 102) = 19.42, p < .001$ ; 年齢,  $F(2, 102) = 4.12, p = .019$ ）。

列挙しないならば、きちんと文法的に成り立っている文にする必要があります。統計部分は必ずカッコに入れます。あくまでも「文」が基本で、統計の数値は添え物です。

- ……3 要因分散分析を行った。その結果、乗り物の主効果は有意で ( $F(2, 204) = 5.67, p = .004$ ), 性別の主効果も有意となり ( $F(1, 102) = 19.42, p < .001$ ), さらに年齢の主効果も有意だった ( $F(2, 102) = 4.12, p = .019$ )。

### 悪い例 (2) 箇条書き

- × ……3 要因分散分析を行った。その結果、  
乗り物の主効果 :  $F(2, 204) = 5.67, p = .004$   
性別の主効果 :  $F(1, 102) = 19.42, p < .001$   
年齢の主効果 :  $F(2, 102) = 4.12, p = .019$   
となった。

統計に限らず、箇条書きは原則としてやめましょう。論文は「文」です。必ず作文して下さい。そして、文は段落になっていなければなりません。

あまりにも検定結果が多くて文にするとわけが分からなくなる場合には、図表にまとめることも検討してください。また、そもそも本当にそんなに大量の検定結果を書く必要があるのか、よく考えてください。論文の本筋とあまり関係のない、重要でない結果ならば、ざっくり割愛するか、付録にまとめてしまうのも手です。教員は、どれだけたくさん検定をしたかでレポートを評価するわけではありません。

### 悪い例 (3) テンプレ文の繰り返し

- × ……3 要因分散分析を行った。その結果、乗り物の主効果は有意だった ( $F(2, 204) = 16.31, p < .001$ )。性別の主効果は有意だった ( $F(1, 102) = 62.62, p < .001$ )。年齢の主効果は有意だった ( $F(2, 102) = 6.68, p = .002$ )。乗り物と性別の交互作用は有意だった ( $F(2, 204) = 11.18, p < .001$ )。乗り物と年齢の交互作用は有意でなかった ( $F(4, 204) = 1.11, p = .352$ )。性別と年齢の交互作用は有意でなかった ( $F(2, 102) = 0.54, p = .586$ )。3 者間交互作用は有意でなかった ( $F(4, 204) = 1.58, p = .182$ )。

これを読んでも、結局どういう結果なのかまるでわかりません。読者に何を理解してもらいたいのかを考えず、ただ検定結果を書き写しているだけです。検定結果は整理して、筋道立てて説明しましょう。ストーリーのある文が大事です。

- ……3 要因分散分析を行った。まず、すべての主効果が有意だった(乗り物,  $F(2, 204) = 16.31, p < .001$ ; 性別,  $F(1, 102) = 62.62, p < .001$ ; 年齢,  $F(2, 102) = 6.68, p = .002$ )。図 2 からわかるように、平均得点は男性より女性の方が、また年齢が高いほど高くなっていった。ただし、乗り物と性別の交互作用が有意だった( $F(2, 204) = 11.18, p < .001$ ) ため、乗り物と性別の効果は単純には解釈できない。この交互作用を反映して、乗り物の単純主効果は男性では有意だったが( $F(2, 204) = 25.67, p < .001$ )、女性では有意ではなかった( $F(2, 204) = 1.81, p = .166$ )。つまり、男性では乗り物の種類による好みの違いが出ていたと言える。これ以外の交互作用は、3 者間交互作用を含めていずれも有意ではなかった( $ps > .10$ )。まとめると、性別にかかわらず年齢が高いほど得点は高くなり、男性では乗り物の種類による違いが見られた。

そこで次に、男性での乗り物の効果を詳しく検討するため、乗り物の単純主効果について下位検定を行うと、……

赤字の文のように、要所要所で検定結果の要点を言葉で説明しなおすだけで、段違いにわかりやすくなります。読者の頭の中にグラフを描いてあげるイメージです。また、あまり重要なポイントでなければ有意でない結果はまとめてしまうなど、検定結果が多いときにはメリハリをつけましょう(ただし、授業レポートでは勉強のため検定結果は省略せずすべて書くように指示されることもありますので、先生の指示に従ってください)。

検定結果の記述では次のポイントを意識しましょう。

1. 要所要所で、つまりどういう結果だったのかを言葉でまとめる。
2. グラフや表を見ながら検定結果を読むとわかりやすいので、図表を適宜参照する。
3. 単に「差があった」だけでなく、差の方向性(どちらが高かった・低かったのか)を言葉で付け加える。

多要因の分散分析で交互作用が出ているなど、検定結果が複雑な場合には、特に注意が必要です。また、あまりに複雑で大量の検定結果がある場合には、図表にするのも手です。

### 3.2 検定結果はあくまでも言葉で説明する

検定結果の書き方というと、どうしても数値や記号の書き方ばかり気にしがちですが、実は一番大切なのは文です。論文ですから、あくまでも文が主で、検定統計量や  $p$  値、グラフや表は従なのです。まず本文があり、補足としてカッコの中に統計量や  $p$  値が添えてある、というのが論文のスタイルです。ですから、何よりもストーリーをはっきりさせて、筋道立てた説明をすることが重要です。

その意味で、「結果」を書くときには次のようなポイントも意識しましょう。

- ・分析や検定の目的を先に書くとわかりやすくなります。「〇〇かどうか確かめるために」「〇〇の効果を検討するために」〇〇検定を行った、というように。
- ・検定の目的をはっきりさせるためにも、検定のデザインを過不足なく言葉で説明すること。特に、対応のある検定をした場合には明記すべし。分散分析の場合は、少しくどいと感じても、要因は何か、各要因は何水準か、各要因は参加者内か参加者間か、などを改めて明記しましょう。また、「 $t$ 検定」「 $\chi^2$ 検定」にもいろいろありますから、どの検定なのかきちんと書きましょう。
- ・結果の説明を図表に丸投げしてはいけません。「反応時間の結果は図2のとおり。」のような1文で済ませるのはダメです。必ず、どういう結果だったのか文で説明し、参考として図表を参照するようにします。
- ・検定の結果も、重要なポイントは言葉で言い換えて説明します。(3.1の「悪い例(3)」参照)。本文を書く前に、検定結果を眺めて、つまりどういう結果だったのか言葉で箇条書きのメモを作ってみましょう。それを整理して、本文にします。

もし結果を言葉でうまく説明できないならば、それは書き手が自分のデータを解釈しきれていないということです。よく頭をひねってください。

それでも、どうしてもやっぱりデータを言葉に置き換えにくいこともあります。中途半端な結果だったり、互いに矛盾するような結果がいくつもあったり、3者間交互作用が複雑すぎてどう考えても理解できない場合など。これはもう、仕方がない。どうしてもわからなかったら、素直にわからないと書けばいいのです。

- この点については解釈が難しいが、総合考察で議論する。
- 予想に反して3者間交互作用が有意だったが、実験の目的とは関連が薄く、効果量も小さいので、解釈しない。

こういうのはOKです。すべての結果を100%鮮やかに説明できる研究なんてめったにないです！ただ、敗北宣言には違いありません。本当にどうしてもわからない時だけにしましょう。

### 3.3 「考察」で何を書けばいいか

論文では、統計的検定結果に基づいて議論するのが原則です。と言うと、検定結果を「結果」で書いてしまったら、「考察」で他に何を書いたらいいのかわからない、と言われることがあります。「考察」は「考察」ですから、推測が入ってもかまわないのです。一生懸命頭をひねって、解釈や推測を繰り返し広げてください。

ただし、ここまでは検定結果から直接言えること、ここから先は推測、とはっきり区別をつけて書くことが大切です。また、たとえ推測であっても、できるだけ状況証拠や部分的な証拠(参加者の内観報告とか、同じような解釈をしている先行研究とか)を挙げるようにしましょう。



### どこにも有意差がなかった場合

「本当に差がない」可能性と、「本当は差があるのに、実験の失敗で差が出なかった」可能性の2つがあります。それぞれの可能性がどれくらいありそうか、考察してください。そして可能ならば、どちらの可能性の方が高いと考えられるか、間接的な証拠も挙げながら結論してください。

後者の可能性はいくらでも考えられます。天井効果や床効果、交絡変数の統制が不十分だった、サンプルサイズが小さすぎた、独立変数の操作が不十分だった、などです。交絡の可能性については、データを何らかの基準で分けて分析してみるとヒントが得られることがあります（たとえば、参加者の性別で分けてみる、成績が高かった群と低かった群に分けてみる、ある方略を使っていたことを実験後に報告していた参加者とそうでない参加者に分けてみる、など）。サンプルサイズは、本来は実験前に先行研究などを参考にして必要な数を推定すべきなのですが、そこまでしていないこともありますし、先行研究より予想外に分散が大きくてサンプルサイズが不十分だったということもあります。検出力（ $1-\beta$ ）や効果量を参考に議論するとよいでしょう。

### 予想外の有意差があった場合

仮説検証型の研究が多い心理学のレポートや卒論では、どうしてもすべて予想通りでないといけない気がしてくるものですが、予想外の結果は何も悪くありません。むしろ、予想外になるからこそ実験の意義があるというものです。予想外の結果を「実はこういうことだったんだ！」とビシッと解釈している論文は、これぞ科学的発見！という感じがしてカッコイイものです。ぜひ、予想外の結果をどうやったら説明できるか、考えてみてください。

とは言え、この場合もやはり「本当に予想が間違っていた」可能性と、「本当は予想が正しかったのに、実験の失敗で別の結果になってしまった」可能性があります。後者の可能性（予想外の交絡変数など）についても、考察をめぐらせてみましょう。

### 論文の本題とはあまり関係がないところに差があった場合

無視せず、ぜひ考察してみてください。たとえば、予想外の男女差が見られた、なんてことはよくあります。研究の目的としては特に男女差には興味がなく、ただ念のため参加者を男女ほぼ同数にただけなんだけれども、結果を見ると実は男女ではっきり違う結果になっていた、というようなケースです。

論文の本筋と関係が薄くても、何か面白いことを示していないか、未知の何かの現象が背後にないかなど、可能性をいろいろと推測してみましょう。新しい研究につながるヒントが見つかるかもしれません。「実は〇〇だった可能性も考えられる。したがって、原因を明らかにするためには、〇〇を統制した上で〇〇を操作するなど、別の検討が必要だろう。」といった調子です。