統計きほんのき

注 1）

心理統計を基本にしているので、医学に応用するには若干不足もしくは過剰なところがあるかもしれませんが、ご容赦ください。

以下の本を参考にしているので、より詳しく学びたいひとはこちらを買ってください。こちらのほうが断然わかりやすいです↓

南風原朝和『心理統計学の基礎』有斐閣アルマ、2002 年

何か質問ありましたら hxfukunaga@gmail.com（福永）までお願いします。

注 2）

基本的に試験の事例問題でわかっておいたほうがよさそうな言葉は赤字もしくは青字にしました。

逆に、読み飛ばしてよいだろうなという箇所はフォントを小さくしてあります。

重要度： 赤字 ＞ 黒字 ＞ 黒小文字

ただ統計がらみの問題はそこまでは出ず、さらに知識問題はほぼ出ていないようなので、何か勉強の際の基礎的な知識の確認にでも使ってください。H19～H24 の過去問で出ているのは、ざっと見たところ

- 調査計画の作成
- 変数の妥当性の確保
- データの偏りの排除
- 検定手続きの選択（帰無仮説の設定から棄却）

のあたりのようなので、そのあたりを意識してさっと読んでいけばいいと思います。

数式の部分は読み飛ばし、研究計画を決める際に必要な知識に絞って憶えていくと良い気がします。

実習で調査するときにはおそらくこれでは全然物足りないので、より詳しい資料をあたってください。また作るかもしれません。

目次

第 1 章　研究法と統計

研究のプロセス

研究は何らかの「問い」すなわち**リサーチ・クエスチョン**（reseach question: RQ）をもつところから始まる．リサーチ・クエスチョンについて考えた暫定的な説明のことを，**仮説**(hypothesis)と呼ぶ．

例：
RQ「思春期における反社会的行動の増加の原因は何か」
仮説「思春期の生理的変化によって生じる心理的葛藤が有力な原因のひとつである」

このような仮説を生み出す過程（**仮説生成**）や，RQ そのものを明確にしていく過程において，行動観察や面接・フィールドワークなどの質的な調査が重要な役割を担っている．
一方，統計的なデータ解析は，生成された仮説の妥当性を検証する（**仮説検証**）過程で主に利用される．

仮説検証型の研究では，データによってその正否が直接的に評価できるような具体的な予測を仮説から論理的に導出する．その予測通りの結果が得られれば仮説は支持され，逆に予測に反する結果が得られれば，仮説は支持されないことになる．
※予測通りの結果が得られたからと言って，仮説が「証明」（逆の場合は「反証」）されるわけではないことに注意．

■研究のアプローチとデザイン

調査（survey）	現実に手を加えず，そのまま把握する
実験（experience）	研究者の側で意図的に条件を**操作・統制**し，その効果を調べる
実践（practice）	対象者の現実の改善を主目的として，介入しながら研究を進める

横断的研究（cross-sectional study）	あるひとつの時点で異なる集団を比較する研究
縦断的研究（longitudinal study）	同一の集団について 2 つ以上の時点を比較する研究

■母集団とサンプル

母集団（population）：RQ や仮説が，少なくとも暗黙のうちに想定している一般的な対象者集団
サンプル（標本）（sample）：実際にデータを収集する対象者集団
被験者（最近は**調査（実験）協力者**，**調査参加者**等と呼ぶ）：データを提供してくれる人．
例：東京都在住の男子高校生 200 人から，友人関係に対する満足度を調査する場合，母集団は例えば「日本の男子高校生」となり，サンプルは実際にデータを取った男子高校生 200 人のことを指す．

■構成概念の測定

構成概念（construct）：学問上の目的のために定義され，使用される概念

構成概念はしばしば抽象的であり，直接的には測定できない．概念の定義を明確にしない限り，測定の方法を定めることができない．構成概念の得点化の仕方に関するルールを，**尺度**（scale）と呼ぶ．そして一定の尺度に従って，被験者に数値を割り振るプロセスを**測定**（measurement）と呼ぶ．

第 2 章　分布の記述的指標とその性質

代表値　：分布全体を 1 つの値で代表するもの

中央値（median）　：分布に含まれる値を大きいものから小さいものへ順に並べたときに順位が全体のちょうど真ん中になる値。データ数が偶数の場合は、中央の 2 つの値の平均を中央値とすることが多い。
「分布に含まれる各値との差の合計が最小となる」値である。

平均（mean）　：データの値の総計をデータ数で割った値。
「分布に含まれる各値との差の二乗の合計が最小となる」値である。

※平均の代表値としての特徴（欠点として言われることが多いが、必ずしもそうではない）
・・・極端に大きかったり小さかったりする値（外れ値 outlier）の影響を受けて変動しやすい
→それらを除いたうえで平均を計算する（**外れ値処理**）を行うことがある。このようにして求めた平均を**調整平均**（trimmed mean）とよぶ。

分布の散布度

散布度（dispersion）：ある変数の分布を要約する統計量。データの範囲（range）・分散・平均偏差・標準偏差などがある。

平均偏差（mean deviation; MD）：各測定値と中央値が平均してどれくらい離れているかを表す値．「中央値からの平均偏差」ともいう．

$$MD = \frac{1}{N}\sum_{i=1}^{N}|x_i - Med|$$

対して，Med の代わりに平均 $\bar{x}$ を入れた

$$MD' = \frac{1}{N}\sum_{i=1}^{N}|x_i - \bar{x}|$$

を使うこともある．こちらは「平均からの平均偏差」というが，これを「平均偏差」と呼ぶこともある．

分散（variance）： $$s^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

標準偏差（standard variation）： $$s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$ SDとも表す.

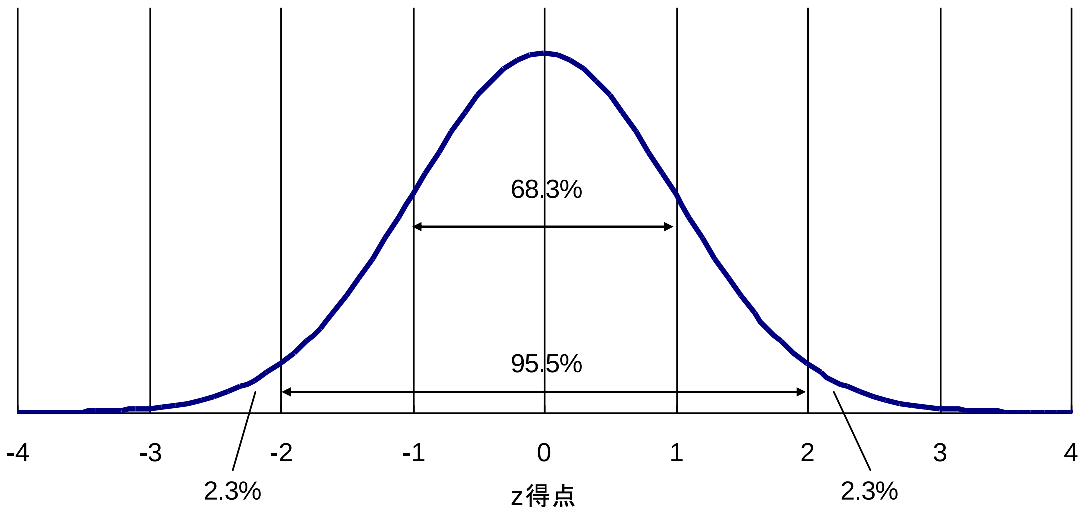
変数の標準化

特定の平均と標準偏差をもつように変換した変数のことを，一般に標準得点（standard score）と呼び，標準得点を求める手続きを**標準化**（standardization）という.（例：偏差値）

その中でも，平均を0，標準偏差を1とする変換によって得られる変数は Z **得点**と呼ばれている。

$$z = \frac{x - \bar{x}}{s_x}$$

標準偏差の解釈←重要！！



変数が正規分布に従うと仮定すると，上記のようにデータが分布する. このことから大まかに，「標準偏差とは，“平均±標準偏差”の範囲に全体の約 3 分の 2 が含まれ，“平均±2×標準偏差”の範囲にはほとんどの人が含まれるような幅のことである」と解釈することができる.

第3章　相関関係の把握と回帰分析

共分散と相関係数

共分散(covariance)：　$s_{xy} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$

不偏共分散：　$s'_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$

s_{xy} のとりうる値の範囲は，　$-s_x s_y \leq s_{xy} \leq s_x s_y$

これをそれぞれ $s_x s_y$ で割って，　$-1 \leq \frac{s_{xy}}{s_x s_y} \leq 1$

この $\frac{s_{xy}}{s_x s_y}$ のことを**相関係数**(correlation coefficient)と呼び，r で表す．

他にも相関係数と呼ばれる指標があるので，それらと区別する際には**ピアソンの積率相関係数**(Pearson's product-moment correlation coefficient)と呼ぶ．

回帰直線のあてはめ

相関関係は，「x の値が大きいほど y の値も大きい(小さい)傾向があるのか」に注目する

回帰分析においては，「x の値の差異に対応して y の値がどの程度異なるか」を考える．

そこで，それぞれの x の値ごとに，それに対応する y の平均に注目する．このような平均を，「x の値を与えたときの y の**条件付き平均**」という．

いま，y を x の 1 次式で予測することを考え，

$\hat{y} = a + bx$　　　（$\hat{y}$: y の**予測値** predicted value。y の条件付き平均を近似する値）

と表す．この式で表される直線を，変数 x から変数 y を予測するときの**回帰直線**(regression line)と呼び，b を**回帰係数**(regression coefficient)と呼ぶ．回帰係数は「変数 x の 1 単位の差異に対応する y の予測値の差異の大きさ」を表し，x の値に対応する y の**条件付き分布**に注目した相関関係の指標といえる．

データに回帰直線をあてはめ，そこから得られる予測値や残差をもとにデータを解釈していく手法を，一般に**回帰分析**(regression analysis)とよぶ．

回帰分析においては，
予測に用いられるほうの変数 x を**独立変数**（or **予測変数** or **説明変数**）と呼び，
予測されるほうの変数 y を**従属変数**（or **目的変数** or **基準変数**）と呼ぶ．

相関係数と回帰係数の性質の違い

相関係数：

対称的（ x と y の相関係数＝ y と x の相関係数）．

標準偏差を用いて相対的に解釈することしかできない．尺度の単位に影響されない．

回帰係数：

非対称的（ y の x への回帰直線≠ x の y への回帰直線）．

用いられた変数の尺度をそのまま用いて具体的に解釈できる．尺度の単位の影響を受ける．

決定係数：証明は省略。従属変数の分散に対する予測値の分散の割合は r^2 である。このことを、「従属変数の分散のうち、独立変数で説明できる割合は r^2 である」と表現できる．このことから， r^2 を分散説明率（proportion of variance accounted for）もしくは決定係数（coefficient of determination）と呼ぶ．

◆選抜効果

何らかの選抜によって調査や実験の対象となる集団が等質化されたとき，その選抜の前後を比較すると，相関係数の値は一般に低下する．この現象を**選抜効果**と呼ぶ．

一方，回帰係数においては， x における選抜を行う限りにおいては，大きな影響を受けない．

このように，相関係数は人為的な要因によって変化しやすいので，複数の研究結果を比較する場合などは，相関係数よりも回帰係数を報告するほうが有用であるといえる．

相関と共変と因果

◆変数間の関係のタイプ

（1） x が大きい人ほど， y も大きい．　→集団における相関関係

（2） x が大きくなると， y も大きくなる．　→個人内の共変関係

（3） x を大きくすると， y も大きくなる．　→処理–効果関係

（4） x が大きいから， y も大きい．　→因果関係

・異なるタイプ間の推論を行き来することはできない．

・相関関係や回帰直線によって直接調べられるのは，(1)のみ．

測定の妥当性と信頼性

測定の妥当性や信頼性を評価する際には，様々な変数間の相関係数が中心的な役割を果たす．

測定の**妥当性**(validity)：測定値が測定すべき構成概念を正しく反映している程度．

測定の**信頼性**(reliability)：測定値が，異なる測定時期や異なるバージョンで一貫している程度．

☆　妥当性があるならば、信頼性は必ずある。逆に、信頼性があるからといって妥当性があるとは限らない。

◆妥当性が相関係数に与える影響

測定値の妥当性が低い場合，変数間の相関係数が低くなるケースと高くなるケースがある．

・低くなるケース：

反映すべき構成概念以外の要因によって変数の値が大きく左右されてしまうために，変数間の相関係数が低くなる．

・高くなるケース：

対象となる 2 つの変数が同じ攪乱(かくらん)要因を共有している場合，2 つの変数が同一の攪乱要因によって連動するので相関係数が高くなる．

◆信頼性が相関係数に与える影響

測定値の信頼性が完全でない程度に応じて，測定値間の相関係数は真値間の相関係数より低くなる．この現象を，**相関の希薄化**(attenuation)と呼ぶ．

第 4 章　確率モデルと標本分布

統計的推測の基本的な考え方

サンプル(標本)(sample)：

調査や実験の際にデータを提供する元となる被験者集団．または，そこから得られたデータの集合．サンプルに含まれる被験者の数を**サンプルサイズ**または**標本の大きさ**という．

母集団(population)：

研究仮説において想定している集団．または，そこから得られたデータの集合．母集団の成員の数は**母集団の大きさ**という．

サンプリング(標本抽出)(sampling)：

母集団からその一部をサンプルとして取り出すこと．

無作為抽出(simple random sampling)：

母集団からランダムにサンプリングすること．ランダムサンプリングの仮定によって，**データの独立性**(個々のデータがどういう値を取るかが，他のデータの値やその確率に全く影響しないこと)が保証される．

標本統計量(sample statistic):

サンプルにおける記述的指標の値. 単に**統計量**ともいう. 標本平均, 標本相関係数など.

母数(parameter):

母集団における記述的指標の値. 英語のまま**パラメタ**ともいう. 母集団平均, 母集団相関係数など.

☆統計的推測の課題は,「**標本統計量**の値をもとに**母数**について出来るだけ正確な推測をすること」である.

★一般的に、標本統計量には英字を、母数には対応するギリシャ文字を割り当てる。

例えば標本平均は M, 母集団平均は μ といった具合。

確率モデル:

「母集団からのサンプリングによって, 1 個 1 個のデータがどういう確率でどういう値を取るか」を表す概念や理論.

確率分布:

確率変数がどういう確率でどういう値を取るかを示す分布.

標本分布:標本統計量の確率分布.(「特定の標本におけるデータの度数分布」ではない)

確率分布における平均・標準偏差の性質

一般に標本統計量は, その標準偏差が大きいほど, その統計量に基づく母数の推定の誤差が大きくなる可能性が高い. その意味で, 標本統計量の標準偏差は, その統計量の**標準誤差**(standard error)と呼ばれる.

N(サンプル数)が増えるほど、標本統計量の標準誤差は小さくなる. 従って, N を大きくすれば, 標準誤差を小さくし, 統計量が母数に近い値を取る可能性を高めることができる. このような性質はどの記述統計量にも共通する重要な性質である.

正規分布モデルと平均の標本分布

心理学の研究では母集団分布に**正規分布**(normal distribution)を仮定することが常套(とう)手段となっており, 実際にほとんどの場合, 正規分布がモデルとして仮定されている. その理由としては,

- 心理学の研究で取り扱う量的変数の中には, その内容から正規分布が最適なモデルとなったり, 経験的に実際のデータの度数分布からみて, **母集団分布に正規分布を仮定することが自然な変数が少なくないから**.
- 正規分布を仮定したほうが**種々の統計量の標本分布の数学的導出が相対的に容易であり, 扱いやすい**ことが多いから.

というのがある.

正規分布に従う確率変数は理論的には**連続変数**(continuous variable)であり, その変数が特定の値をとる確率ではなく, ある特定の範囲の値を取る確率を問題にする. データ分布が正規分布と有意差があるかどうかを検定する方法として、コルゴモロフ-スミルノフ(Kolgomorov

Smirnov)検定がある。正規分布においては，平均と標準偏差を示せば分布が完全に決まる．

■平均の標本分布

一般的に，N 個のデータで構成されるサンプルの平均 $\bar{x}$ については，母集団分布の種類に関係なく，$\bar{x}$ の標本分布の平均が母集団平均 μ に等しくなることが知られている．したがって，$\bar{x}$ の標本分布の平均が母集団平均の不偏推定量（標本統計量の分布の平均が，その統計量によって推定しようとしている母数の値に一致するような標本統計量）となる．

確率モデルの適用に関する諸問題

■母集団の設定

身近な学校や病院などに協力依頼をして集めたサンプルの場合，それに対応する母集団がどういう集団なのかがはっきりせず，結果を過度に一般化してしまう危険性がある．これを防ぐためには，先行研究の蓄積や，研究者の持つ知識などに基づいて，一般化可能な範囲を主観的に判断し，限定していく必要がある．

■サンプリングのランダム性

実際のサンプルは母集団からランダムに選ばれたものではないため，確率モデルを適用するための手続き的根拠を欠いている．

これについては，

① 明確に定義された母集団からのランダムサンプルでない限り，確率モデルを前提とした統計的推測を禁止する立場と，

② 対象となる変数の内容や統計量の種類を考慮して母集団を限定したら，実際のサンプルをその母集団からのランダムサンプルとみなす立場がある．

①の立場ならば，確率モデルの適用による誤りを完全に排除することができるが，サンプルサイズの違いに基づいてデータの優劣を判断することができないため，小さなサンプルから得られた統計量の値に過度の信頼を置く危険性がある．
②の立場ならば，確率モデルの適用により種々の統計的推測が可能だが，適用がどれくらい正当化できるか明らかでないのがデメリットとしてある．

■ランダム化

実験研究においては，ランダム化（無作為化，確率化）によってデータに確率的な変動の要素を入れることで，ランダムサンプルでないデータに確率的な視点を導入する根拠とすることがある．

しかしこの手続きも被験者集団内部での変動であり，データ集合としての母集団についての議論とは直接結びつかない．やはり，結果をどのような母集団に対して一般化できるかについては，その都度統計学の枠を超えたやり方で判断していかざるを得ない．

また，調査研究にはランダム化は行えないため，確率的視点導入の根拠を与えることができない．

■頑健性

確率モデルとしての母集団分布を設定することが容認されたとしても，どのような確率分布をモデルに採用するかはまた別の問題として残る．

多値の量的変数については正規分布を仮定することが多いが，「もし本当の母集団分布が正規分布でなかったとしたら，正規分布を仮定した結果はどれぐらい妥当なのか」ということが問題になる．このように，仮定したモデルが真の状態と異なる場合に，それでもそのモデルに基づく結果が妥当である程度を，その方法の**頑健性**（robustness）と呼ぶ．

・ 一般に，N が大きいときには母集団分布の違いの影響は小さくなり，頑健性が増す．
・ N が小さかったり母集団分布が非常に偏っている場合には，特定の分布形を仮定しないノンパラメトリック法という方法を用いることもある．

第 5 章　推定と検定の考え方

推定（estimation）と**検定**（statistical test）は，ともにデータから母数に関する推測を行う方法である．

推定には，母数の値はこれぐらいだろうという形で 1 つの値によって母数を推定する**点推定**（point estimation）と，ある区間を設けて母数の値はこの区間に含まれるだろうという形で推定する**区間推定**（interval estimation）がある．

一方，検定では，母数の値に関する仮説を立て，データに基づいてその仮説を採択するか棄却するかの判断を行う．

推定量とその標準偏差

推定量（estimator）：母数の点推定のために用いられる標本統計量
推定値（estimate）：実際のデータから計算される推定量の値
不偏推定量：その期待値が，推定しようとしている母数の値に一致するような推定量

■標準誤差による推定精度の評価

母数の点推定においては，推定量の精度を的確に把握することが重要である．そのために**標本誤差**（母数と推定値の差）を知りたいが，これを正確に把握することは不可能なので，代わりにサンプリングによって推定量が母数のまわりをどの程度変動するかに注目し，推定量の標本分布の標準偏差，すなわち**標準誤差**に注目することによって推定量の精度を把握する．

この標準誤差を一定以下に抑えるという視点から，必要なサンプルサイズを計算，決定することもできる．

検定の考え方

2 変数正規分布である母集団分布に従うデータにおける，相関係数の検定を例に考える．

相関係数に関する検定のうち最も単純かつ典型的なものは，「母集団相関係数はゼロである」という仮説を立て，データに基づいてその仮説が棄却できるかどうかを判定する，いわゆる無相関仮説の検定である．検定される仮説は多くの場合，それを棄却する目的で設定されるので，一般に，「無に帰するべき仮説」という意味で**帰無仮説**(null hypothesis)と呼ばれている．帰無仮説の検定に用いられる統計量を一般に**検定統計量**(test statistic)と呼び，帰無仮説のもとでの検定統計量の分布を**帰無分布**(null distribution)と呼ぶ．

棄却域(rejection region)：帰無仮説と整合的でないとされる検定統計量の値の範囲．

棄却の限界値：棄却域の端点．

統計的に有意(statistically significant)**である**：

確率的に偶然とは考えにくく，意味があると考えられるという意味．検定統計量の値が棄却域に入り，帰無仮説が棄却されるとき，その検定統計量は「**統計的に有意**である」または単に「有意である」という．

統計的に有意であるということは，実質的に意味があるということを意味するわけではなく，単に帰無仮説が正しいとしたら得られにくい値というだけの意味である．

両側検定：帰無分布の両側に棄却域を設定する検定方式．

片側検定：帰無分布の片側に棄却域を設定する検定方式．

有意水準(significant level)：

「帰無仮説のもとではまれにしか生じない事象」というものを定義する確率の値．α で表される．

棄却域を定めるために設定される基準である．

p **値**：

得られた結果が有意となる有意水準の最小値．**限界水準**，**有意確率**とも呼ばれる．データから計算される統計量である．

検出力とその応用

検出力(power)：母集団においてゼロでない(相関や平均値差などの)母数があるとき，サンプルにおいて帰無仮説の棄却の際に有意な結果が得られる確率．帰無仮説が偽のとき，それを棄却できる確率．

第 1 種の誤り(type I error)：帰無仮説が正しいときに，それを棄却してしまう誤り．

第 2 種の誤り(type II error)：帰無仮説が正しくないときに，それを採択してしまう誤り．

第 1 種の誤りの確率は有意水準 α と同じ.
第 2 種の誤りの確率は β と表記する.
検定力は $1-\beta$.

N を大きくすると，検定力が高まる．一定の検定力を確保するのに必要なサンプルサイズを決める手続きを，**検定力分析**（power analysis）という.

区間推定の考え方

「あらかじめ定められた確率で母数を含む区間」を考える．この，「あらかじめ定められた確率」のことを**信頼水準**（confidence level）または**信頼係数**と呼び，その確率で母数を含むと推定される区間を**信頼区間**（confidence interval）と呼ぶ.

例えば，信頼水準 .95 で母集団相関係数を含むと推定される区間は，母集団相関係数の 95% 信頼区間と呼ぶ.

区間推定では、
「得られた標本統計量から，母数に関する帰無仮説を棄却できないような母数の範囲」
を求める.

区間推定においても， 点推定における推定量の標準誤差の場合と同様に，N が大きいほど幅は小さくなる．そこで，予想される相関係数の値と，望まれる信頼区間の幅からサンプルサイズを決めることができる.

第 6 章　平均値差と連関に関する推測

独立な 2 群の平均値差の検定

■前提

2 つの群の間で従属変数 y の平均を比較する統計的推測では，y の母集団分布としてそれぞれの群ごとに正規分布を仮定し，**各群の母集団分布の分散が互いに等しい**という仮定をおく（SPSS のプログラムでは「等分散性の検定」として出てくる！）.

また，**各群のデータは互いに独立である**と仮定する．（この仮定が成り立たないのが，**対応のある**群の場合）

帰無仮説（H_0 と表記する）は、$H_0: \mu_1 = \mu_2$（両群の母集団平均が等しい）

■平均値差の標本分布

平均値差 $\mu_1 - \mu_2$ に関する検定や推定には，標本平均の差 $\bar{y}_1 - \bar{y}_2$ が用いられる．各群のデータが互いに独立な場合には，その差 $\bar{y}_1 - \bar{y}_2$ の標本分布も正規分布となることが知られている．

$H_0 : \mu_1 = \mu_2$ が正しいとき，標本平均値差 $\bar{y}_1 - \bar{y}_2$ をその標準誤差の推定量 $s_{\bar{y}_1 - \bar{y}_2}$ で割った比

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s_{\bar{y}_1 - \bar{y}_2}} \qquad (6.3)$$

は，t 分布に従う．そこで，t 分布の上側確率値を用いることによって，平均値差を正確に検定することができる．上式の t を検定統計量とし，その値を t 分布に照らして行う平均値差の検定を，**t 検定**（t test）もしくは独立な 2 群の t **検定**という．

■検定力とサンプルサイズ

2 群の平均値差の検定における検定力は，

- 母集団における効果量 $\delta = \dfrac{\mu_1 - \mu_2}{\sigma}$
- サンプルサイズ $\underline{N}$，$\underline{n_1}$，$\underline{n_2}$

によって規定される．

対応のある 2 群の平均値差の検定と推定

マッチング：被験者の対を作ること．

ブロック：マッチングによって作られる被験者の対．

対応のある群には，

被験者の対または組が作られるケースと，

個々の被験者自身がブロックとなるケースがある．

帰無仮説は、$H_0 : \mu_1 = \mu_2$

■平均値差の標本分布

ブロック内のデータの差を　$v = y_1 - y_2$　とすると，

$$\bar{v} = \bar{y}_1 - \bar{y}_2$$

となる．対応のあるデータ y_1，y_2 の間に高い正の相関があるほど，対内の差 v の標準偏差 σ_v が小さくなり，その結果 $\bar{v}$ の標準誤差 $\sigma_{\bar{v}}$ が小さくなる．

2 群の比率の差の検定

■独立な 2 群の比率差の標本分布

質的変数に関する比率の群間差 $p_1 - p_2$ に注目する．

比率 p の標本分布は，サンプルサイズ N を試行数，母集団比率 π を成功確率とする 2 項分布によって与えられる．この分布については，中心極限定理により，N がある程度以上大きければ，正規分布で十分正確に近似できることが分かっている．

■独立な 2 群の比率差の検定

$H_0 : \pi_1 = \pi_2$（両群の母集団比率が等しい）

■対応のある 2 群の比率差の検定（カイ二乗検定ではない）

$H_0 : \pi_1 = \pi_2 \quad \Leftrightarrow \quad n_{12} = n_{21}$

2 つの母集団比率が等しいときの，$n_{12}(n_{21})$ の確率分布を考え，それを標本の $n_{12}(n_{21})$ と照らし合わせる．そして標本の $n_{12}(n_{21})$ が棄却域に入るかどうかを検討する．

例） 化学と物理における合格/不合格のクロス集計表

$n_{12} + n_{21} = m$ とおき，m を固定して考えると，H_0 の下での $n_{12}(n_{21})$ の分布は，試行数 m，成功確率.5 の 2 項分布に従う．

	物理合格	物理不合格	合計
化学合格	n_{11}	n_{12}	$n_{1\cdot}$
化学不合格	n_{21}	n_{22}	$n_{2\cdot}$
合計	$n_{\cdot 1}$	$n_{\cdot 2}$	N

→方法①

この分布の両端もしくは一端に棄却域を設定し，それに $n_{12}(n_{21})$ が入るかどうか調べる．

→方法②

$\mu_w = N\pi = m/2$，$\sigma_w = \sqrt{N\pi(1-\pi)} = \dfrac{\sqrt{m}}{2}$ の正規分布に近似し，検定統計量を標準化して標準正規分布を参照する．

カテゴリ変数間の連関の分析

一方の項目においてある選択肢を選んだ被験者のほうが，そうでなかった被験者より，もう一方の項目において特定の選択肢を選ぶ傾向があるのかといった，2 つのカテゴリ変数の間の関係のことを**連関**（association）とよぶ．　*cf.* 相関

カイ 2 乗統計量 χ^2 は，2 つのカテゴリ変数の間に連関が全くない状況を想定し，その状況下

で期待される結果と実際のデータとの差異を評価するものである.

いま，変数 x と y がそれぞれ a 個と b 個のカテゴリを持つとし，x のカテゴリを表側に，y のカテゴリを表頭に配置した連関表を作成し，i（i=1,2,･･･，a）行目 j（j =1,2,･･･，b）列目のセルの度数を n_{ij} とする．全体の度数を N，x の周辺度数を $n_{i\cdot}$．y の周辺度数を $n_{\cdot j}$ とする．

2 変数間に連関が全くない場合の（i , j）セルの度数の期待を e_{ij} とすると，

$$e_{ij} = N \times \frac{n_{i\cdot}}{N} \times \frac{n_{\cdot j}}{N} = \frac{n_{i\cdot} n_{\cdot j}}{N}$$

となる．この値を，連関がないときの推定期待度数とよぶ．

カイ 2 乗統計量 χ^2 は，e_{ij} と n_{ij} との差異を次式で評価したものである．

$$\chi^2 = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{(e_{ij} - n_{ij})^2}{e_{ij}}$$

例） 化学と物理における合格/不合格のクロス集計表（実際の標本）

	物理合格	物理不合格	合計
化学合格	n_{11}	n_{12}	$n_{1\cdot}$
化学不合格	n_{21}	n_{22}	$n_{2\cdot}$
合計	$n_{\cdot 1}$	$n_{\cdot 2}$	N

例） 連関がないと仮定したときの化学と物理における合格/不合格のクロス集計表

	物理合格	物理不合格	合計
化学合格	e_{11}	e_{12}	$e_{1\cdot}$
化学不合格	e_{21}	e_{22}	$e_{2\cdot}$
合計	$e_{\cdot 1}$	$e_{\cdot 2}$	N

■連関の決定

母集団において 2 つのカテゴリ変数間の連関が全くないという帰無仮説の下で，カイ 2 乗分布に χ^2 が近似的に従うことを利用．自由度 $(a-1)(b-1)$ と有意水準の値によって決まるカイ 2 乗分布の一定の上側確率の値を，データから計算された χ^2 の値が超えれば，帰無仮説が棄却され，統計的に有意な連関があったということになる．

■連関の検定と比率差の検定の関係

2 群の比率差の検定のための統計量 z と，2×2 の連関表のためのカイ 2 乗統計量 χ^2 の間には，

$$\chi^2 = z^2$$

という単純な関係があり，z を用いた両側検定とカイ 2 乗検定とは常に同じ結果になる．

第 7 章　実験デザインと分散分析

実験デザインと要因

分散分析(analysis of variance; ANOVA)：

質的な独立変数の値によって従属変数の平均がどのように異なるかを分析するための方法．

独立変数の水準が 3 つ以上あるときには t 検定を用いることはできないので、この手法を用いる．

基本的な帰無仮説は、「各群の母集団平均に差がない」！検定統計量は F！

要因(factor)：質的な独立変数．

水準（level）：質的な独立変数の値．

心理学の実験においてどの要因をとりあげるか決める場合，従属変数に影響を与える可能性のある要因は無数に考えられる．現実的に考えると研究の焦点をその中の少数の要因に絞らなければならないが，このとき問題になるのが研究で取り上げなかった残りの要因の処理である．これらをうまく処理しないと，2 つ以上の要因が連動して変化し，そのうちのどれが結果に影響したかわからなくなる．この状態を，要因が**交絡**(confound)しているという．交絡を避けるには，とりあげない要因を**統制**(control)する必要がある．

◎統制の方法

一定化：　研究でとりあげる要因のどの水準においても，統制する要因の水準が同じになるように条件を調整すること．

バランス化：　実験でとりあげる要因のどの水準にも，統制する要因における各水準の被験者が同数ずつ割り当てられるようにすること．

ランダム化（**無作為化**）：　実験でとりあげる要因の各水準に，被験者をランダムに割り当てること．

対応のない要因：異なる水準に含まれる従属変数の値が，互いに独立である要因．

対応のある要因：異なる水準に含まれる従属変数の値に相関がある要因．

・・・被験者内要因　→　対応のある要因

・・・被験者間要因　→　各水準にランダムに割り当てれば対応のない要因。
マッチングを行うなどすれば対応のある要因

◆　対応づけることの意味

- 実験でとりあげないある要因における特定の水準の被験者が，実験でとりあげる要因のある水準に集中するのを避けることができる．
- 要因の効果に対する検定力が高くなる．

◆　実験デザインの種類

要因の数，対応の有無，被験者間か被験者内かによって区別される．

対応のない 1 要因デザイン（完全無作為 1 要因デザイン）

対応のない 2 要因デザイン（完全無作為 2 要因デザイン）

対応のある 1 要因デザイン（1 要因ランダムブロックデザイン，1 要因反復測定デザイン）

などなど．

完全無作為 1 要因デザイン

◆ 仮定と検定統計量

「各水準における従属変数の母集団分布が，一定の分散を持つ正規分布である」という仮定

$$y_{ij} \sim N(\mu_j,\ \sigma_\varepsilon^2)$$

をおく．この仮定の下に，帰無仮説

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_a$$ （全ての水準の母平均が等しい）

をおき，これが正しいとき，F 統計量が F 分布に従うことを利用して，検定を行う

多重比較の考え方

全体として有意な結果が得られた後に，どの水準とどの水準の間に有意差があるか調べる検定を，一般に**事後検定**(post hoc test)という．

複数の群の比較（検定）を行うことを**多重比較**(multiple comparisons)という．

多重比較に伴う統計的な問題とは，

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_a$$

が正しいときに，それを誤って棄却する確率，すなわち第 1 種の確率の問題である．この帰無仮説を，対ごとの検定を繰り返すことによって検定するとしたら，対ごとの検定を 5%水準で行った場合，上記の帰無仮説に対する第 1 種の誤りの確率は 5%より大きな値となってしまう．これを防ぐため，テューキーの方法などの特別な検定法が多重比較のための方法として提案されている．

完全無作為 2 要因デザイン（おまけ）

一方の要因の効果のあり方，すなわち水準間の平均値差が，他方の要因の水準によって異なるとき，「2 つの要因の間に**交互作用**(interaction)がある」という．

交互作用効果に対し，それぞれの要因ごとの効果のことを**主効果**(main effect)という．

検定により、主効果や交互作用があるかどうかを推測する．

仮定： 各セルにおける従属変数の母集団分布が一定の分散 σ_ε^2 をもつ正規分布である

帰無仮説： 要因 A の主効果，要因 B の主効果，及び要因 A ×要因 B の交互作用効果がゼロである

対応のある 1 要因デザイン

「ブロック」を要因として扱い，2 要因デザインとして処理する．詳細は SPSS がやってくれてるので、これも試験では知らなくてもよいと思われる。

◆ 仮定と検定統計量：仮定を省略。検定統計量は F 統計量。
  帰無仮説は、「要因 A の効果がゼロである」

共分散分析

　完全無作為デザインでは残差となる個人差を，回帰分析のように量的な変数によって予測し，それによって，モデルで説明できない残差部分を小さくすることができる。