実験計画学
2007 年 11 月 20 日

第６回　統計的検定

Ａ．統計的検定

１．統計的検定とは？

統計的検定：統計的に標本の統計量から母集団の母数に関する予測の真偽を検証することを統計的検定という．

統計的検定の例：

例１　さいころを５回振ったところ，５回とも奇数だった．二項分布から５回とも奇数になる確率は 1/32（=0.03125）である．このような低い確率が出ることから，さいころは奇数と偶数が同じ確率で出ると考えるよりも，奇数が出やすいと結論した.

例２　同じM寸の卵でありながらスーパーＳは，スーパーＫよりも軽い卵を売っていると考え，両店の 10 個の卵についてそれぞれ調査したところ, スーパーＳの卵は平均 20g, 標準偏差 0.5g, スーパーＫの卵は平均 22g，標準偏差 0.5g であった．したがって，スーパーＳの卵の重さは信頼区間 95％で 20±0.31g であり, スーパーＫの卵の重さは信頼区間 95％で 22±0.31g である．２つの信頼区間は重ならないこの結果から両店の卵の重さの母平均は違うと結論づけた.

２．検定の手順

①　帰無仮説の設定

統計的検定によって否定したい仮説を立てる．これを**帰無仮説**という．例１ではさいころは奇数と偶数が同じ確率（0.5）で出る，例２ではスーパー２店の卵の重さが同じであるというのが帰無仮説に相当する．そして，帰無仮説に従えば，実際に得られたデータの出現する確率はいくらかを求め，その確率がきわめて低い（基準は必要によって定める）ならば，帰無仮説が間違っているとし，対立仮説が正しいとするのが統計的検定である.

②　対立仮説の設定

帰無仮説が間違っていると判断した場合に採用する仮説を**対立仮説**という.

では，対立仮説は例１，２ではどうなるか？

例１：

例２：

③　帰無仮説が成り立つとしたら，今回得たデータが出現する確率を求める．
この確率のことを**有意確率（p−値）**と呼ぶ．

④　p−値があまりに小さいと，帰無仮説が正しいと考えるより間違いであると判断する．これを「**帰無仮説を棄却する**」という

★　一般には p−値が 0.05（5%），0.01（１%），0.001（0.1%）などより小さいときに，帰無仮説が間違いであると判断する．

⑤　帰無仮説が成り立たないと判断したときには対立仮説を採用する．

有意差がある：帰無仮説を棄却できるだけの反証が得られた．

p−値が 0.05（あるいは 0.01or0.001）より小さいときに，5（あるいは 1 or 0.1）%の**有意水準**で**有意差**があると表現する．

⑥　帰無仮説を棄却できないときは，有意差がないことになり，対立仮説を採用するだけの証拠がないことになる．

帰無仮説が棄却できないときは，帰無仮説を採用するのではなく，帰無仮説を棄却できるだけの証拠が不十分と考える．この場合，採用するのは**対立仮説の否定**である．

対立仮説を証明したいのに，帰無仮説のようなものを持ち出すのはまどろっこしく感じるかもしれない．しかし，統計的にできるのは帰無仮説の否定だけである．その理由は以下の２つである．

★　「甲は乙より早い」といっても，どのくらい速いのか，甲が乙に勝つ確率はいくらなのか，がはっきりしない．これを明確に規定できなければ，数量的に取り扱えない．
★　仮に，仮説を数量的に規定できても，それが真実であることを統計的には証明できない．統計的にできることは，仮説に対する「反証」を提示することだけである．

Ｂ．二項分布を利用した統計的検定

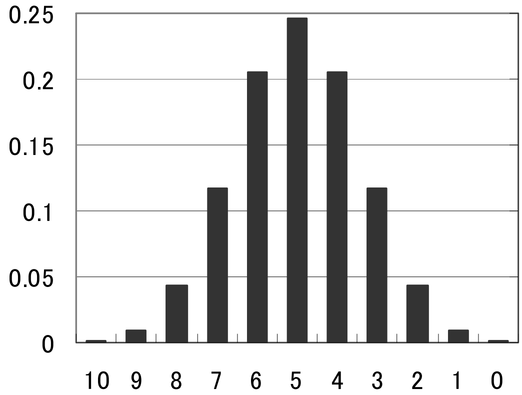
１．二項分布を利用した統計的検定

例題：Ａ君は対戦型ゲームでキャラクターＢと互角に戦えると主張しているが，実際やってみると１勝９敗だった．Ａ君の主張の真偽を統計的に検定せよ．

帰無仮説：Ａ君はキャラクターＢと互角に戦う．
　　　　　Ａ君がキャラクターＢに勝つ確率は（　　　　）である．

対立仮説：Ａ君はキャラクターＢより弱い（互角でない）．
　　　　　Ａ君がキャラクターＢに勝つ確率は（　　　　）ではない．

Ａ君がキャラクターＢと互角に戦う（勝つ確率 0.5）としたら，二項分布に従って，Ａ君が１勝以上しない確率を計算する



Ａ君がちょうど１勝する確率

Ａ君がちょうど０勝する確率

1 勝だけの確率をみるのではなく，どれだけ中心から外れているかの確率を計算するので，0 勝の場合も計算しなければならない

２．有意水準と検定結果

① 有意水準

仮説が正しいにもかかわらず，帰無仮説を棄却する確率を（　　　　　　　）という．
（　　　　　　　）には 5%，1%，0.1%などがよく使われる．

ここでは有意水準を 5%としてみよう（検定前，p-値を計算する前に有意水準を決めること）．

帰無仮説が正しいとすると今回のようなケース（Ａ君が１勝あるいは０勝しかしないということ）が起こる確率（有意確率，p-値）は（　　　　）であった．

帰無仮説は棄却できるだろうか？

★　5%の有意水準で帰無仮説は（　　　　　　　），（　　　　　　　　　）と判断できる．

★　1%の有意水準で帰無仮説は（　　　　　　　），（　　　　　　　　　）と判断できる．

② 帰無仮説を検定する

帰無仮説が棄却される→対立仮説を採用する

帰無仮説が棄却されない→対立仮説の採用には証拠が不十分であり，この場合，対立仮説の否定的な表現となる（「対立仮説」とはいえないという表現をする）

練習：あるチョコレートには 0.2 の割合で当たりくじが入っているとメーカーは主張する．10 個買ったところ，1 つも当たりくじが入っていなかった．メーカーの主張の真偽を有意水準 5%で検定せよ．

帰無仮説：

対立仮説：

p-値＝

検定結果

検定結果：

有意水準 5%で帰無仮説は（棄却される・棄却できない）．

したがって，

1）対立仮説を採用する→

2）対立仮説は採用できない→

３．第 1 種の誤りと第 2 種の誤り（有意水準と検出力）

① 第 1 種の誤り

帰無仮説が正しいにもかかわらず，これを否定する過誤を第 1 種の誤りという．例 1 では本当はさいころは正しくできているのに，たまたま 5 回続けて奇数が出たなら，第 1 種の誤りを犯すだろう．

第 1 種の誤りを犯す危険率を**有意水準**といい，α と表すことがある．有意水準を小さくして，検定すれば，第 1 種の誤りを犯す危険は小さくなる．さいころの例でいえば，10 回続けて奇数が出なければ，帰無仮説を否定しないとすれば，有意水準 0.001 で検定したことになり，そういうことは 1000 回に 1 回しかないので，帰無仮説を否定できると結論できる．

② 第２種の誤り

しかし，5 回続けて奇数の出るさいころが正しいさいころであるよりは，奇数の出やすいさいころと考える方がありそうである．このように帰無仮説が誤りであるにもかかわらず，これを否定しない過誤を第２種の誤りという．第１種の誤りを減らすために有意水準を小さくすれば，第２種の誤りの危険率（β）は高くなり，すなわち**検出力**（$1-\beta$）は低下する．２種類の誤りと検定結果には下の表のような関係がある．

表 検定における２種類の誤り

		本当に成り立っているのは	
		帰無仮説	対立仮説
検定結果	帰無仮説	正しい（その確率：$1-\alpha$）	第２種の誤り（その確率：β）
	対立仮説	第１種の誤り（その確率：α）	正しい（その確率：$1-\beta$）

有意水準を小さくすれば，第１種の誤りをおかす危険性は（大きく・小さく）なる．しかし，第２種の誤りをおかす危険性は（大きく・小さく）なる．

統計的検定では有意水準を決めることによって，第（　　）種の誤りの大きさを制御している．すなわち第（　　）種の誤りを避けることに重点が置かれている．このことは統計的検定では貴重な発見を見落としてしまってでも疑わしい結果を受け入れることを避けているのである．

C．母分散が既知のときあるいは大標本の平均に関する統計的検定

１．ある決まった平均に対する検定

① 両側検定（通常の検定）

標本平均は標本の大きさが十分大きければ正規分布に近似できる．そこで，正規分布の性質から母集団から得られた標本の平均がある値と異なるかを検定することができる．

例：A農場で出荷する桃の重さの標準偏差σは 20g である．桃の重さの平均が 150g となるように出荷する．ある日の調査では 100 個の桃の重さを量ったところ，平均 145g であった．桃の重さの母平均は 150g ではないのかを検定せよ．

帰無仮説 H_0：

対立仮説 H_1：

① 帰無仮説が成り立つとすると

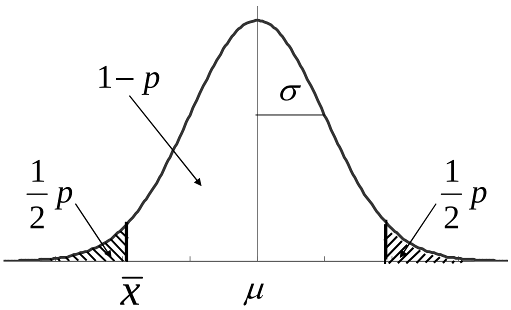
母平均は（　　　　）g であり，母標準偏差は（　　　　　）g である．標本数は（　　　　　）個であるから，標準誤差は（　　　　　）g である．したがって，標本平均は平均（　　　　　）g，標準偏差（　　　　　）g の正規分布に従う．

② 今回のデータが得られる確率を計算する

今回の標本平均は（　　　　　）g であるから，この標本平均が得られる確率 p−値はエクセルの NORMSDIST 関数で計算できる．

得られた標本の標本平均を $\overline{x}$，帰無仮説でとりあげた母平均を μ，標準誤差を SE とすると，今回の標本が帰無仮説の下で得られる確率（p−値）は以下のように計算できる．

p−値 $= 2 \times NORMSDIST(-ABS(\overline{x} - \mu)/SE)$



帰無仮説 $\mu = 150$ が成り立つとき，
今回の標本 $\overline{x} = 145,\ n = 100$ が得られる確率
p−値＝（　　　　　　　　　　　　）である．

したがって，有意水準 5%では帰無仮説 H_0 は棄却される（有意水準 5%で有意である）．すなわち対立仮説 H_1 を採用し，有意水準 5%でこの標本の母平均は 150g ではないと結論できる．

しかし有意水準 1%では帰無仮説 H_0 は棄却されない（有意水準 1%で有意でない）．すなわち対立仮説は採用できないので，有意水準 1%でこの標本の母平均は 150g ではないとはいえないと結論される．

練習：C 村は塩分の濃い食事で有名であり，寿命が短いといわれる．村民の平均寿命は 100 人調べたところ，70.2 歳，標準偏差は 0.9 歳だった．平均寿命が 70 歳であるかを有意水準 5%で検定せよ．

帰無仮説　H_0：

対立仮説　H_1：

p−値＝

検定結果

② 片側検定 どちらか一方に変化したことが予想できる場合

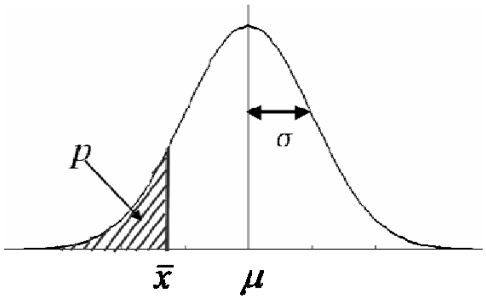
例：Ｂ牧場では牛のえさをＦ社からＧ社に変えた．Ｇ社のえさはより栄養価が高いので，Ｆ社のときの泌乳量 5.0L，標準偏差 0.8L より向上すると予想した．実際に 100 頭を調査した結果，泌乳量は 5.2L となった．標準偏差は変わらない．泌乳量（の平均）が増加したかを検定せよ．

帰無仮説 H_0：

対立仮説 H_1：

p−値＝両側検定の p−値÷2

p−値＝ $NORMSDIST(-ABS(\bar{x}-\mu)/SE)$



帰無仮説 $\mu = 5.0$ が成り立つとき，

今回の標本 $\bar{x} = 5.2,\ n = 100$ が得られる確率

p−値＝（　　　　　　　　　　　）である．

したがって，有意水準 1%で帰無仮説H_0は棄却される（有意水準 1%で有意である）．すなわち対立仮説H_1を採用し，有意水準 1%でこの標本の母平均は 5.0L よりも増加したと結論できる．

このようにある一方向に対する検定（片側検定）では有意になりやすくなる．しかし，これはあらかじめ<u>泌乳量が増加するという情報がある</u>からである．たいていの場合，こういう片側検定を使うことは望ましくない．少なくとも実験を始める前にどういう検定をするかを決めた場合でないと使わない方がよい．

２．２つの独立した母集団から得た２つの標本平均についての検定

それぞれ無作為に抽出した２つの標本について，２つの母平均が異なっているかを検定する．

例：イネの１穂穎花数について 100 本ずつ調査し，品種Ａは 150±12，品種Ｂは 155±16 を得た（平均±標準偏差）．両品種の１穂穎花数（の平均）は異なるのかを検定せよ．

この場合，２つの標本の平均の差が 0 であるという帰無仮説を検定する．２つの標本の平均の差も正規分布に従う．ここで２つの標本から得た標準偏差を一緒にする必要がある．２つの独立した母集団では１つの母集団にすると分散は両分散の和になるという性質がある（詳しい証明は略する）．このことから標本数 n_A の標本（標本平均 $\bar{x}_A$，標準偏差 σ_A）と標本数 n_B の標本（標本平均 $\bar{x}_B$，標準偏差 σ_B の２つの独立した標本を得たとき，次のことがいえる．

$\bar{x}_A - \bar{x}_B$ の分布は正規分布で，$N(\mu_A - \mu_B, \frac{{\sigma_A}^2}{n_A} + \frac{{\sigma_B}^2}{n_B})$ になる．

帰無仮説　H$_0$：

対立仮説　H$_1$：

p−値$= 2 \times NORMSDIST(-ABS((\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)) / \sqrt{\frac{{\sigma_A}^2}{n_A} + \frac{{\sigma_B}^2}{n_B}})$

$$= 2 \times NORMSDIST(-ABS((\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)) / SQRT(\frac{\sigma_A \wedge 2}{n_A} + \frac{\sigma_B \wedge 2}{n_B}))$$

計算結果

したがって，有意水準 5%で帰無仮説を棄却できる（両側検定）．すなわち２つの品種の 1 穂穎花数（の平均）は異なると結論できる．

練習：20～24 歳男子 2112 人の身長は 170.5±5.9cm，30～39 歳男子 2776 人の身長は 169.5±5.8cm（人間生活工学研究センター1992-1994 年調査）であった．２つの母集団の間に平均身長は差があるかを検定せよ．

帰無仮説　H$_0$：

対立仮説　H$_1$：

p−値＝

検定結果

このようにサンプル数が十分あれば標準誤差が小さくなるので，平均に関する検定における検出力は高くなる．個々の標本については逆転することも多々ある（標準偏差の大きさからもわかるように分布はかなり重なる）が，母平均に有意差があることは証明できる．

D．統計的検定の手順のまとめ

１．帰無仮説をたてる．

２．対立仮説をたてる．ふつうは両側検定とする．

３．有意水準を決める．

有意水準は統計的検定の計算を始める前に決める．可能ならば実験を始めたり，データを集めたりする前に決めるのが望ましい．とくに統計によって何らかの判断を検定結果から行うときは有意水準を事前に決めなければならない．p−値をみてから有意水準を決め，判断するというのでは，統計的な方法をつかわないで恣意的にカンで物事を決めるのとほとんど同じことである．

統計による判断を実験した人が行わないで，ただデータをつけて結果の提示だけを行うときにはデータに p−値あるいは有意差を示す記号を付ける．p−値をつけるのがより明確だが，以前は p−値の計算が面倒であったので，慣例として有意差を示す記号を付け，*は 5%水準で，**は 1%水準で，***は 0.1%水準で有意差があることを示すことが一般的である．しかし，この記号は慣例で使われているだけなので，確認して使うべきである．

４．p−値を計算する．

特に有意水準について断りのない場合は，得られた p-値について以下のように結論するのが一般的である．

0.05＜p-値　5%の有意水準で帰無仮説は棄却できず，対立仮説の否定を採用する．
0.01＜p-値≦0.05　5%の有意水準で帰無仮説は棄却され，対立仮説を採用する．
0.001＜p-値≦0.01　1%の有意水準で帰無仮説は棄却され，対立仮説を採用する．
p-値≦0.001　0.1%の有意水準で帰無仮説は棄却され，対立仮説を採用する．

５．統計的検定を行う．

p−値が最初に決めた有意水準よりも小さければ，帰無仮説は棄却され，有意差があると判断できる．帰無仮説が棄却できないときは「差があるとはいえない」という結論になる．帰無仮説を採用して，「差がない」という結論にしてはいけない．

６．現実的な意味での差があるかの判断

統計的に有意差があることと現実的な意味でその差に意味があるかというのは別問題である．すなわち p−値が小さければ小さいほど，差があることを統計的に強く確信が持てるとはいえる．しかし p−値が小さければ小さいほど，現実的な意味での価値が増すというわけではない．

E．宿題

1．第５回の宿題２．で調査したデータについて帰無仮説は２人の間に実力差はないとし，有意水準（危険率）を 5%として，二項分布するという仮定の下で検定せよ．なお授業では片側検定として検定したが，ここでは両側検定を行うこと．

2．第３回の宿題で調べたデータについて，今回返した前回の宿題の講評の最後に書いてあることについて検定を行え．

3．太平洋に浮かぶ２つの島（Ｘ島，Ｙ島）に共通する鳥であるＰ鳥は毎年，たくさんの卵を産む．しかし，Ｙ島の近くに重油タンカーが座礁したため，卵の重さがＸ島と異なるかもしれない．両島で卵をそれぞれ無作為に 100 個ずつ採集したところ，Ｘ島では 35.0±1.2g，Ｙ島では 34.6±1.1g の値（平均±標準偏差）を得た．両島のＰ鳥の卵の重さの母平均は異なるのかを 1%の有意水準で両側検定せよ．

4．２つの別々のスーパーから 10 個入りの卵（M寸）を買い，両店の卵の重さは同じなのかを調べる．ここでは，卵の重さをはかり，平均と標準偏差を計算する．ここで得た平均と標準偏差から正規分布となる母集団のグラフをだいたいで書いてみる．さらに 95%信頼区間をつけた母集団の平均の推定値をそれぞれについて求める．そこから両店の卵の重さが同じかを考えて見よ．(同じようなデータであれば，卵でなくてもよい．下の例を見よ．２人一組で班をつくって，調査してもかまわない)

有効数字が 3 つ以上になるようにデータをとること．すなわち 28.1g，1.53cm のようにデータをとること．

例：２つの店のラーメンの重さ
２人のソフトボールの遠投距離
２人の 50m 走

補足　統計的検定の結果の表記について

１．有意水準をかならず付けて表記すること

p−値が 0.05 より小さく，0.01 より大きい場合，５％の有意水準では有意であり，１％の有意水準では有意ではない，すなわち５％の有意水準では２つの標本の母平均に差がある，１％の有意水準では差があるとはいえないという結論になる．このときときどき差があるといいながら，あるとはいえないともいうのかと混乱する人もいるかもしれない．あくまでも宣言した有意水準において差がある，差があるとはいえないと結論しているのであり，統計的検定においてはかならず有意水準を付けなければならない．

検定の結果を次の行動につなげる場合はかならず有意水準を決めてから検定しなければならない．この場合，有意水準α％で検定せよということになり，もし p−値が p%であれば，結論は以下のようになる．

p≦α　有意水準α％で有意差である．有意水準α％で母平均に差がある．
p＞α　有意水準α％で有意差がない．有意水準α％で母平均に差があるとはいえない．

とくに有意水準を指定しない場合は以下のように表記するのが一般的である．

p＞5%　有意水準５％で有意差がない．有意水準５％で母平均に差があるとはいえない．
1%＜p≦5%　有意水準５％で有意差がある．有意水準５％で母平均に差がある．
0.1%＜p≦1%　有意水準１％で有意差がある．有意水準１％で母平均に差がある．
p≦0.1%　有意水準 0.1%で有意差がある．有意水準 0.１％で母平均に差がある．

有意水準５％，１％，0.1%で有意差があるときに*，**，***をデータにつけることがある．

２．p−値が小さいことと処理の効果が大きいことは別である

ときどきより小さい有意水準で有意である，すなわち p−値が小さい方が効果が大きいと勘違いする人がいる．しかし p−値が小さいことは，２つの標本の母平均が異なることをより強く確証できるだけであり，効果そのものの大きさとは関係がない．例えば，効果の強い薬を３人程度で比較試験してもかならずしも有意差を得られるとは限らない．すなわち p−値はあまり小さくはならない．一方，効果のそれほど強くない薬でも数万人も使って，試験すれば，かなり小さな p−値を得るかもしれない．

このように p−値の大きさは標本の大きさに大きく依存する．

さらに有意差がなくてもその薬に効果がないということを証明したわけではない．あくまでも帰無仮説が棄却されないということは，差があるとはいえないと結論しただけである．もし３人程度の実験で有意差を得られない薬でも１０人程度で実験すれば有意差が得られるかもしれない．

このように p−値はあくまでも統計的に２つの標本の母平均に差があるかどうかの判断基準を確率によって与えるだけであり，その２つの母平均の差が（現実的な意味で）大きい差か小さい差かは p−値からは判断できない．