

検定

佐賀大学 新井康平

1

定義と例題

- 検定は「最初に仮説を立てて、実際に起こった結果を確率的に検証し、結論を導く」という手順で行います。「はい/いいえ」を用います。例題とは「最初に仮説を設定し、仮説が正しいと判断する条件を考えて矛盾が生じた場合に仮説が間違っている」と判断する」方法
- 50%の確率で表が出る普通のコインと、10%の確率でしか表が出ない不正なコインがどちらか2枚ずつあるとします。偽物この4枚のコインは全く見た目が同じで見分けることはできません。
- 「普通のコインだよ」と言われ、普通の不正のどちらかのコイン2枚組を渡されたとして、渡された2枚のコインを投げる計行を2回繰り返したところ、2回ともすべて裏でした。渡された2枚のコインは普通のコインでしょうか、それとも不正なコインでしょうか。

2

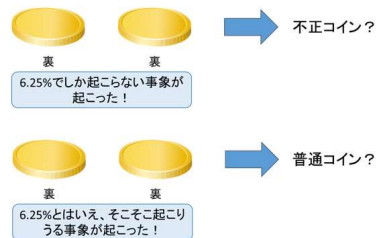
仮説検定

- 仮説を立てる
- 検定を行うにあたっては、まず「仮説」を立てます。ここでは「渡された2枚のコインは普通のコインである」とします。普通の仮説に矛盾が生じた場合、普通のコインではないと結論付けられます。
- 仮説を検定するための「確率」を求める
- 渡されたコインが普通のコインであった場合、「2枚とも裏」が2回起こる確率

$$\left\{ \left(\frac{1}{2} \right)^2 \right\} = \frac{1}{16} = 0.0625 = 6.25\%$$
- これは、もし普通のコインを渡されていたら、6.25%でたまたま起こらない事象が起こったことを示しています。したがって、6.25%という確率をもとに判断すると、次の2つの結論のどちらかになります。
 - 6.25%というのは減少に起こらない確率だ。つまり渡されたコインは不正なコインに違いない。
 - 6.25%というのは低い確率だが、それでも偶然起こることもある確率だ。つまり普通のコインに違いない。

3

仮説検定



4

仮説検定

- 結論を確定させるために「減少に起こらない」と判断する確率の基準」を設定します。例えばこの基準を「10%」と設定すると、10%以下の確率で起こる事象については「減少にないこと」であるため、「仮説が間違っている」と判断されます。逆に、10%以上の確率で起こる事象については、「減少にないことではない＝許容できる範囲」となるため、「仮説が間違っているとは言えない」と判断されます。
- この問題では、仮説を「渡されたコインは普通のコインである」とし、観測した事象が起こる確率は6.25%でした。6.25%は判断基準10%より小さいため「渡されたコインは普通のコインであるとは考えられない」という結論を導くことができます。

5

帰無仮説と対立仮説

帰無仮説

検定の最初に立てる仮説のこと
この仮説を元に検定を行い結論を導く
 H_0 と書かれることがある

対立仮説

帰無仮説に対する仮説のこと
本来証明したい仮説
 H_1 と書かれることがある

6

P値

- 検定統計量から算出されたP値を元に検定の結論を導きます。
- 検定統計量：帰無仮説が正しいと仮定したときに、観測した事象よりも稀なことが起こる確率を計算するための値です。簡単に「統計量」とよばれる場合もあります。
- P値：帰無仮説が正しいとした仮定とき、観測した事象よりも極端なことが起こる確率のことです。「観測した事象よりも極端な事象が起こる確率」であることから、これは累積確率となっています。コインの問題では、「6.25%」がP値になります。
- P値はその大小を比較するものではありません。あくまでも事前に設定した有意水準と比較するためのものです。

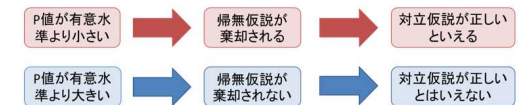
7

有意水準

- 帰無仮説を棄却する基準を有意水準といいます。
- 有意水準：帰無仮説を棄却するための基準となる確率です。この数値は検定を行う前に決めておく必要があります。危険率と呼ばれることもあります。コインの問題では、「減少にない」と判断する確率の基準」がこの有意水準です。
- 棄却：仮説を捨てることです。P値が有意水準よりも小さい時は、帰無仮説を捨て対立仮説を採択します。すなわち、対立仮説が正しいと結論付けられます。コインの問題では、基準である「10%」と観測した事象の起こる確率「6.25%」を比べた結果、「6.25%」が「10%」より小さいことが分かりました。このとき、帰無仮説「渡されたコインは普通のものである」を「棄却」し、対立仮説「渡されたのは不正なコインである」を採択します。

8

仮説採択、棄却



9

有意水準

- 有意水準は、検定において帰無仮説を設定したときにその帰無仮説を棄却する基準となる確率のことです。(アルファ)で表され、**5% (0.05)** や **1% (0.01)** といった値がよく使われます。有意水準は検定を行う前に設定しておきます。
- 有意水準を0.05に設定するということは、「5%以下の確率で起こる事象は、100回に5回以下しか起こらない事象だ。したがってこのようなまれな事象が起こった場合、偶然起こったものではないとしてしまおう」という意味です。したがって、**P値が0.05 (5%) を下回った場合、そのP値は偶然取る値ではないと結論付けられます。**言い換えると、「極めて珍しいことが起こった」あるいは「何かしら意味があることである (= "有意である") 」ということを表します。

10

検出力

有意水準

α で表される
帰無仮説を棄却する基準
第1種の過誤(帰無仮説が正しいのに棄却する)を犯す確率

検出力

$1-\beta$ で表される
帰無仮説が正しくないときに、正しく帰無仮説を棄却する確率
 β は**第2種の過誤**(対立仮説が正しいのに採択されない)を犯す確率を示す

11

		真実	
		帰無仮説が正しい	対立仮説が正しい
検定の結果	帰無仮説を棄却しない (対立仮説が正しいとは言えない)	正しい	第2種の過誤() 0
	帰無仮説を棄却する (対立仮説が正しい)	第1種の過誤() 0	正しい() 0

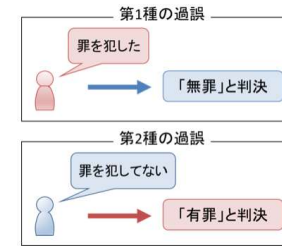
12

		真実	
		本当は有罪 (帰無仮説が正しい)	本当は無罪 (対立仮説が正しい)
裁判の結果	有罪と判決 (対立仮説が正しいとは言えない)	正しい	第2種の過誤() 0
	無罪と判決 (対立仮説が正しい)	第1種の過誤() 0	正しい() 0

13

2つの過誤

- ある容疑者の裁判の例では、第1種の過誤を犯す確率を下げるために容疑者を全て有罪にしてしまうと、無罪の人まで有罪になってしまうために第2種の過誤を犯す確率が上がってしまいます。逆に第2種の過誤を犯す確率を下げるために容疑者を全て無罪にしてしまうと、本来有罪の人まで無罪になってしまうために第1種の過誤を犯す確率が上がってしまいます



14

z値とt値

- 平均が0、分散が1となるようにデータを標準化した値

$$z = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

- サンプルサイズがnの場合、統計量tは自由度n-1のt分布に従います。そのため、統計量tを用いた検定を行う際には自由度n-1のt分布を使います。統計量tを用いた検定のことを「t検定」といいます。t検定は、調べる値の母集団が正規分布することが前提条件

$$t = \frac{\bar{x} - \mu}{\sqrt{\frac{s^2}{n}}}$$

15

例題

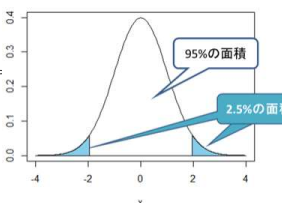
- 日本人の男性100人をランダムに選び、その身長を測定したところ**平均175、不偏分散100**となりました。身長分布は正規分布に従うとする時、**日本人の男性の平均身長は180cm**と断言してよいでしょうか。
- この場合の**帰無仮説は「日本人の男性の平均身長は180cmである」、対立仮説は「日本人の男性の平均身長は180cmではない」となります。**この例題では母分散は分からないので、標本から得られた**不偏分散を使って統計量t**を求めます。

$$t = \frac{\bar{x} - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{175 - 180}{\sqrt{\frac{100}{100}}} = -5$$

16

95%信頼度

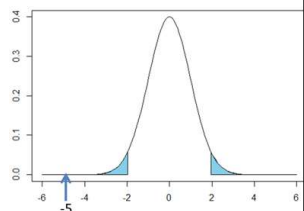
- この統計量tを用いて検定を行います。有意水準95%で検定する時、統計量が次の図のt分布の水色部分に入る場合に帰無仮説は棄却されます。両端の水色部分の面積は合わせると全体の5%であり、統計量tがこの部分に入るということは5%以下でしか起こらない極めて珍しい事象であると判定されます。



17

統計量 t

- この例題では**統計量t=-5**となり、この値は上図の左側の水色部分に含まれるため、有意水準5%では帰無仮説は棄却され、対立仮説が採択されます。つまり、「日本人の男性の平均身長は180cmではない」と結論付けられます。



18

T検定

- Excelでt検定をするには **T.TEST関数**を利用すると簡単に計算できます。T検定は母集団が正規分布に従うと仮定して検定をします。(母集団が正規分布に従っていない場合は、標本平均はt値からは多かれ少なかれ乖離します。)
- ある新薬を開発しました。投薬前と投薬後に測定をしました。結果は下記の表です。このとき、新薬は(数値を下げる)効果があると言えるでしょうか。
- 帰無仮説は、差がない、誤差のない仮説になるため「**投薬前と投薬後に効果の差は無い**」となります。

19

F検定：F.TEST

- 「ペンギン組」(8名)と「アヒル組」(5名)であるテストを実施しました。テストの得点は以下になりました。
- ペンギン組とアヒル組の得点の分散に差があるか調べます。
帰無仮説は誤差が無い仮説になるため、「**ペンギン組とアヒル組の得点分布には差がない**」となります。
- E2セルの値は "0.045..." となります。片側確率は半分の "0.0225..." になります。有意水準を5%とすると、0.0225 < 0.05 となり2つのグループで分散の有意差はあると判断でき、仮説は棄却され、ペンギン組とアヒル組とは分散が違うことが言えます。

20

χ^2 検定：CHISQ.TEST

- おみくじの箱にくじが5つ入っています。4つがはずれで、1つがあたりです。6人がくじをそれぞれ一人5回引き、当たった回数、はずれた回数を記録します。結果は以下となりました。
- このとき、**人によってくじのあたり方に差があるでしょうか**(エスパーみたいな人がいる、または、くじ箱に細工がしてあるなど)くじは公平で誤差の範囲内でしょうか？
- 帰無仮説は差がない、誤差のない仮説となるため「**人によってくじのあたり方に差はない**」となります。
- H2セルの値は、"0.119..." となります。(約12%) 有意水準を、5%(普通用いる統計学的有意水準)としても5%より高い値のため、「人によってくじのあたり方に差はない」の仮説は棄却できず、くじのあたりの出方は人によって変わらないことが言えます。

21