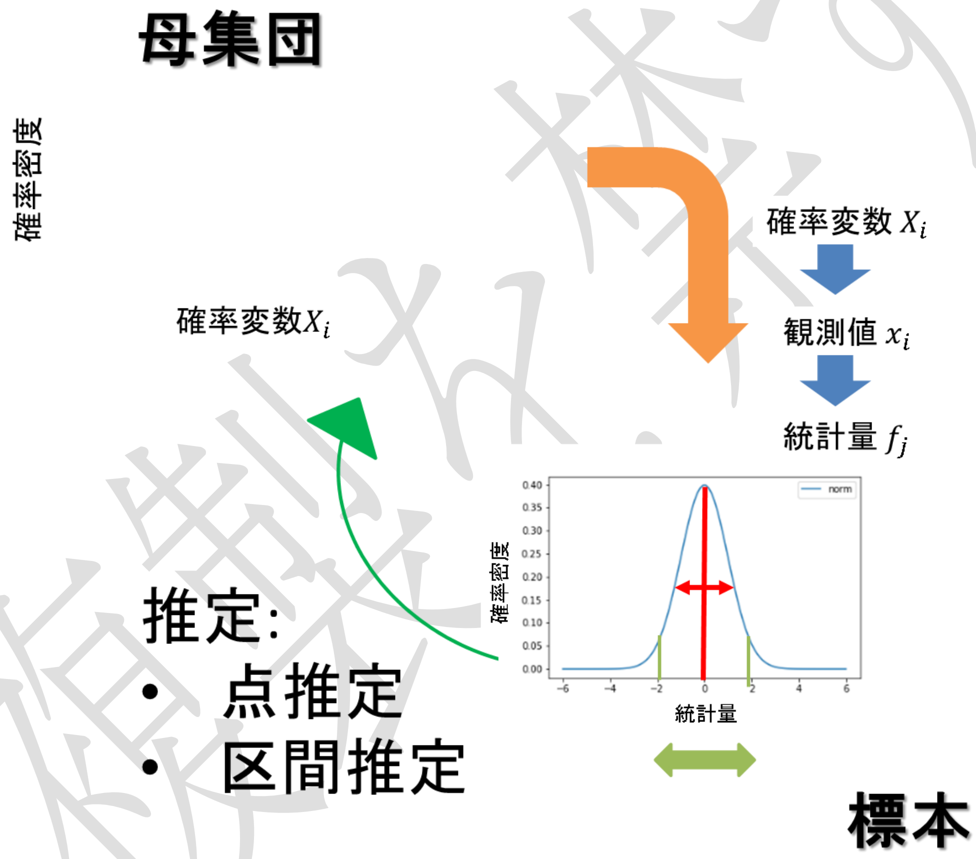
推測統計のはじめは統計的推定です。確率変数と確率分布、そして観測値を基礎とする推測統計を学びます。ここでは、未知の母数を、観測値(得られたデータ)をもとに推測していきます。これを統計的推定の問題といいます。得られたデータ x から未知の母数を考えるとき、**その統計量の推定値とともに、その信頼度も考える必要があります。**つまり、推定値がどの程度の範囲にあるかを考える必要があるのです。母集団からデータ $x_1, x_2, \ldots, x_n$ が得られるとすると母数の推定値は何度も計算することができ、かつその値はいつでも同じではありません。それらを確率変数ととらえるとき、推定量となります。

母数の推定値を表現する方法には２つあり、１つの値としてとらえるのが点推定、上限、下限の間の区間としてとらえるのが区間推定です。



4.1 点推定

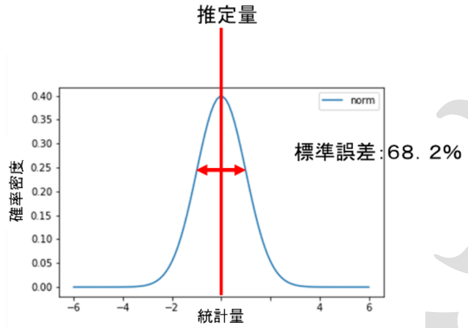
標本 $x_1, x_2, \ldots, x_n$ から算出される１つの値で、未知の母数を推定する方法を点推定といいます。
－ 平均、分散など

母数 θ に対してその推定量は θ に ＾ をつけて表します。

4.1.2 標準誤差

母集団から得られた標本から統計量を推定するとき、そのばらつきの度合いを標準誤差といいます。これは標本のすべての組み合わせの標準偏差で表します。単に標準誤差といったときには平均のばらつきを表し、それは分散の推定量を標本の大きさで割り、その平方根をとったものです。推定量と標準誤差は組として示されます。

- **点推定**
  - **推定量** $\hat{\theta}$
  - **標準誤差(推定量の標準誤差)** $\mathrm{se}(\hat{\theta})$



4.2 区間推定

標本から得られる統計量の上限と下限の 2 つの値を求めて、その間に母数がふくまれるという表現の方法が区間推定です。

- 信頼区間

確率変数を X、区間の上限を$U(X)$、下限を$L(X)$、そして、母数を M とすると、

$$L(X) \leq M \leq U(X)$$

と表現します。M は$U(X)$と$L(X)$の間に入ることを意味します。

- 信頼係数

この信頼区間の中に母数が入る確率が信頼係数で $1-\alpha$ で表します。したがって、

$$P[L(X) \leq M \leq U(X)] = 1 - \alpha$$

となります。$L(X)$、 $U(X)$の決め方が統計的推定に大きな影響を与えます。

4.2.1 母平均の区間推定

母平均の区間推定を行ってみましょう。標本$x_1, x_2, \cdots, x_n$, は独立に平均 μ, 分散 σ^2 の正規分布にしたがうとします。

- 母分散が既知の場合
- 標準正規分布を用います。

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$N(0,1)$は標準正規分布

- z_{α}： 確率 α における標準正規分布の臨界値

1 － α は信頼係数で標本平均が信頼区間に入る確率になります。

$$\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- 標本x_1，x_2，⋯，x_n，は独立に平均 μ， 不偏分散s^2の正規分布にしたがうとします。

母分散は未知です。

- t 分布を用います。

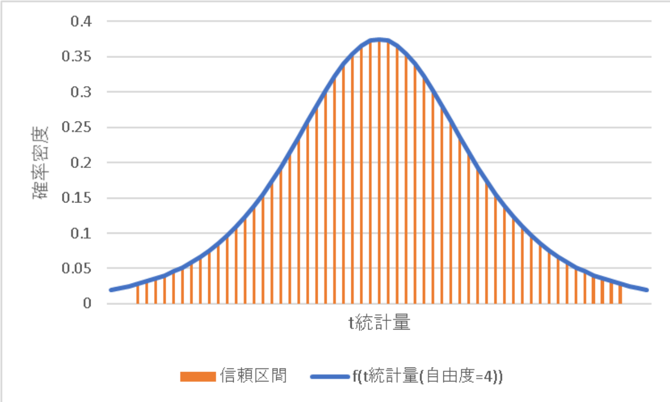
$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim T(n-1)$$

- $t_{(\alpha,n-1)}$： 確率 α 、自由度 n － 1 の t 分布の臨界値

$$\bar{X} - t_{(0.5\alpha,n-1)}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{(0.5\alpha,n-1)}\frac{s}{\sqrt{n}}$$

例題 4.1： ニキビの治療を受けに病院を訪れた５名の患者さんにそれぞれ A,B,C,D,E とローマ字を割り当てます。訪問時の患者 A のニキビの数は１１、B は９、C は１２、D は８、E は１０とします。その際の母平均の推定値を求めてみましょう。信頼係数は９５％とします。

患者のニキビの平均個数は１０個です。不偏分散は２、t 統計量の臨海値は±2.13 です。よって下限は 8.65、上限は 11.34 です。



4.2.2 母分散の区間推定

分散の区間推定をする場合には、カイ二乗分布を用います。標本分散では

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 \sim \sigma^2 \chi^2_{(n-1)}$$

の関係があります。これを変形して、

$$\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

とします。そうすると z はカイ二乗分布にしたがいます。信頼係数 $1-\alpha$ の信頼区間は

$$\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\chi^2_{(0.5\alpha,n-1)}} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\chi^2_{(1-0.5\alpha,n-1)}}$$

となります。また

$$\frac{s^2(n-1)}{\chi^2_{(0.5\alpha,n-1)}} < \sigma^2 < \frac{s^2(n-1)}{\chi^2_{(1-0.5\alpha,n-1)}}$$

標本の大きさが大きくなると信頼区間は狭くなります。

例題 4.1：エクセルによるワインデータの主な要素の区間推定をしてみましょう。

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	平均	8.32	0.53	0.27	2.54	0.09	15.88	46.47	1.00	3.31	0.66	10.42	5.64
2	分散	3.03	0.03	0.04	1.99	0.00	109.42	1082.14	0.00	0.02	0.03	1.14	0.65
3	標準偏差	1.74	0.18	0.19	1.41	0.05	10.46	32.90	0.00	0.15	0.17	1.07	0.81
11	自由度	1598	1598	1598	1598	1598	1598	1598	1598	1598	1598	1598	1598
12	信頼係数	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
13	上限	11.19	0.82	0.59	4.86	0.16	33.09	100.61	1.00	3.57	0.94	12.18	6.97
14	下限	5.45	0.23	-0.05	0.22	0.01	-1.34	-7.67	0.99	3.06	0.38	8.67	4.31
15		A	B	C	D	E	F	G	H	I	J	K	評価
16		7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	5

練習問題 4.1： excel で正規乱数を発生させ基本統計量をとってみましょう。乱数の数を 10、100、1000、10000 といろいろと変えてやってみましょう。

練習問題 4.2： 赤ワインデータベースの一変量要約統計量の最大値、最小値と推測統計から得られる上限と下限を比較してその特徴を述べてみましょう。

練習問題 4.3：ひずんだ分布を修正する方法があるかどうかを試してみましょう。

第５章　統計的仮説検定

推測統計の手法の２つ目は、統計的仮説検定です。本章では前章の推測統計の手法を利用して統計的仮説検定を考えていきます。

成功体験は自信をもたらす

5.1 仮説検定をなぜ統計的に行う必要があるのか？

実際に得られたデータを分析してみると、想定していた値とは異なる値が出るようなときがあります。データは実は想定していた特性をもっていないかもしれません。たまたま得られたデータがまれであるのかもしれません。

成功体験が過ちを繰り返す原因

ところで何度も何度も繰り返す観測の中で、確かに全部がおかしければ、それはおかしいという話になります。しかし、確率的に起こる可能性があるのであれば、それはおかしくないかもしれません。そして、それがどの程度の確率で起こればおかしいと考えたらよいのでしょう。そのような基準になる確率を**有意水準**といいます。そしてこのように判断するしかたが**統計的仮説検定**です。客観的に判断したいときに用います。

過ちを繰り返さないための客観性

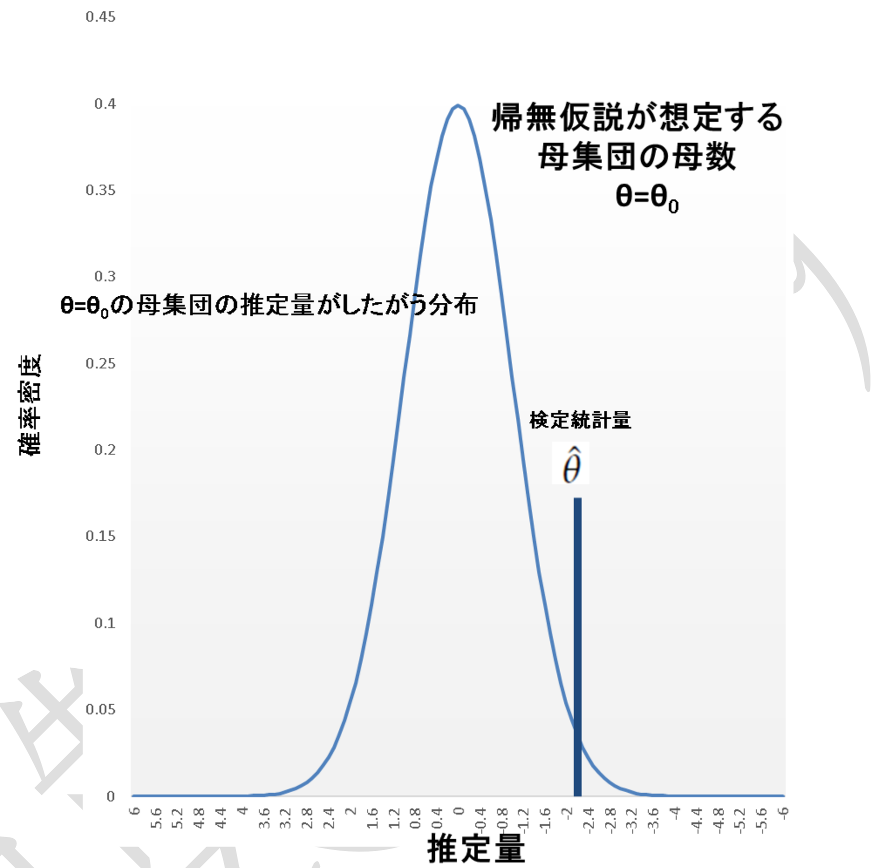
5.2 仮説検定の構造

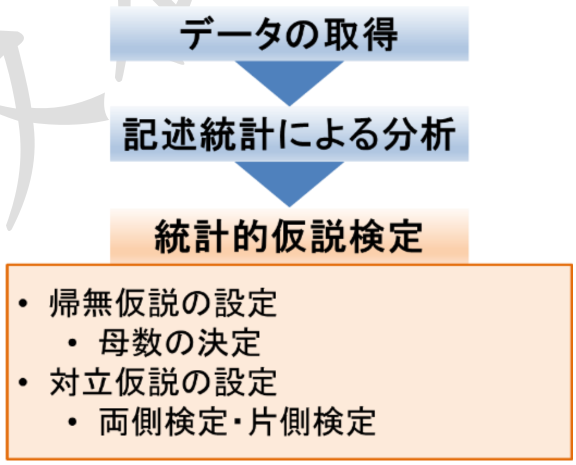
5.2.1 帰無仮説と対立仮説と有意水準

統計的仮説検定では仮説を立てます。その仮説は実際にはすてられてほしい仮説です。それを帰無仮説といいます。帰無仮説では、母集団を定めるのですが、それを母数により行います。そして標本データが”その母集団から得られたものであるかないかを判断し、「ない」“と判断されると帰無仮説は棄却されます。すてられます。帰無仮説が正しくないときに成り立つものを示しているのが対立仮説です。対立仮説には片側検定と両側検定が

あります。

帰無仮説では母数を提示しているので、それに対応する統計量の分布が得られます。観測値から得られる推定値がこの分布からみて稀にしか起こらないと判断されたとき、帰無仮説はすてられます。この判断の基準が有意水準です。これは統計量の分布から見て稀にしか起こらない事象を確率で表したものです。観測値から得られる推定値より稀な事象が起こる確率が有意水準 α より小さければ、帰無仮説に対してかなりまれな事象が起きたとして、帰無仮説はすてられます。そうでなければすてることはしません。帰無仮説をすてたとき、観測値は別の母数で規定される母集団に属すと考えます。このような仮説を統計的仮説とよび、このような検定の方法を、統計的仮説の有意性検定といいます。





5.2.2 帰無仮説

帰無仮説では母数θをθ_0と定め、H_0と書きます。

$$H_0: \ \theta = \ \theta_0$$

帰無仮説の例

- クラゲの占い

ある水族館で飼育されているクラゲが野球の試合の勝敗を予言する能力があるといいます。クラゲはでたらめにチームを選んでいるのかもしれません。それとも予言する能力があるのでしょうか？でたらめに選んでいるのなら成功確率は半々です。したがって帰無仮説は成功確率 p=1/2 とします。

$$H_0: p = 0.5$$

- コインのひずみ

コインのひずみの有無を判断するために、コインを投げることを考えます。ひずみが無ければ表の出る確率と裏の出る確率は等しいので、表の出る確率 p は 1/2 とします。そうすると p = 0.5 が帰無仮説です。

$$H_0: p = 0.5$$

5.2.4 対立仮説

帰無仮説がすてられたときに、どのような仮説がすてられずにすむのかを示すのが対立仮説です。

検定には２つのタイプがあります。1 つは、片側検定です。帰無仮説のときと同様に対立仮説として、すてられずにすむ仮説を母数で表現します。$\theta<\theta_0$はθがθ_0よりも小さいとき、$\theta<\theta_0$は、θがθ_0よりも大きいとき用いられます。

$$H_1: \theta<\theta_0\text{または}\theta>\theta_0$$

もう 1 つは明確に同じであるか無いかを検定したいときです。その場合に、比べたい両者を等しいとおいて、両者がたいへんに異なるときに、おかしいとすればよいのです。その際に、どちらの方向に大きくおかしいかは関係ありません。このような検定を両側検定といいます。

$$H_1: \theta \neq \theta_0$$

対立仮説の例

- クラゲの占い

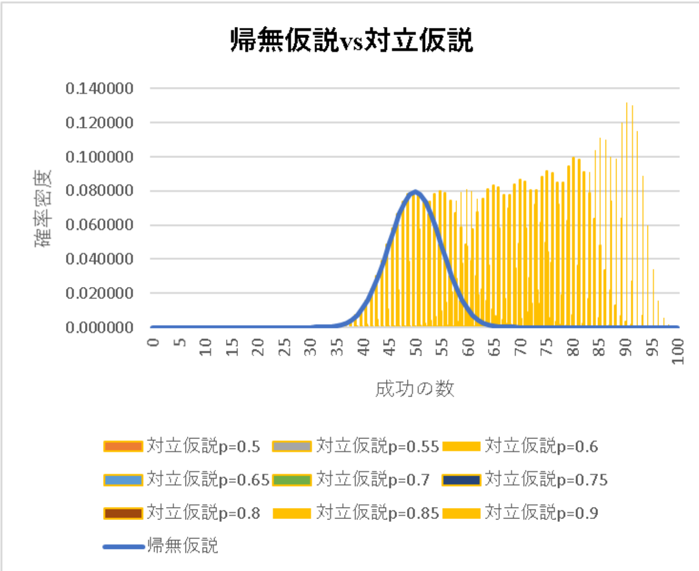
ある水族館で飼育されているクラゲが野球の試合の勝敗を予言する能力があるという場合の帰無仮説は成功確率 p=1/2 でした。この場合でも、成功確率が 2 分の 1 よりも小さければクラゲは勝敗を予言する能力をもたないことになります。成功確率は 2 分の 1 よりも大きくなければなりません。したがって、対立仮説は p>1/2 です。

$$H_1: p>1/2$$

例題 5.1: クラゲの占いの帰無仮説と対立仮説を、分布図を用いて表現してみましょう。試行の回数は１００回とします。

ある水族館のクラゲは野球の試合の勝敗を予測できるといいます。その勝ち負けの選択はクラゲが札を指さすことで実現します。予測が的中すればそれは成功の回数となり、そうでなければ失敗となります。これは二値の選択問題です。したがって、成功の回数の確率は２項分布にしたがいます。その際に成功の確率を p=0.5 とすれば、クラゲはランダムに札を選んでいることになり、予言が的中する確率は五分五分になります。p>0.5

よりも大きければ予測能力があることになります。帰無仮説は青い線で表されています。試行回数１００,p=0.5 の２項分布です。試行回数１００の内もっとも起きる確率の高い成功の回数は５０です。５０は平均ですのでこれを中心にベル型の分布をしています。対立仮説は p=0.5 から 0.05 刻みで分布図を描いていますが、実際には p>0.5 の連続したものです。p=0.6 のときの成功の平均回数は６０回です。p=0.9 のときは９０回です。



グラフはそれぞれの対立仮説が等確率で起きると仮定されています。このように帰無仮説と対立仮説を可視化してみると、帰無仮説は母数をもちますが、対立仮説も母数をもちます。帰無仮説が棄却されれば対立仮説の母数が帰無仮説の母数になる可能性があります。このように見てみると母数にも空間があります。それが母数空間です。

例題 5.2：ニキビに悩む人のニキビの平均の数(母平均)は 10 個であることが分かっています。病院を訪れた患者 A,B,C,D,E の５名に薬を投与したところ、それぞれ９，９，１０，９，８になりました。薬を投与した後のニキビの平均の数は、薬を投与しない一般の人よりも少ないといえるでしょうか？母分散は既知で$\sigma^2 = 2$とします。

ニキビの平均の数を μ とすると

帰無仮説 H_0: $\mu = 10$

対立仮説 H_1: $\mu < 10$

ニキビの数の平均=9

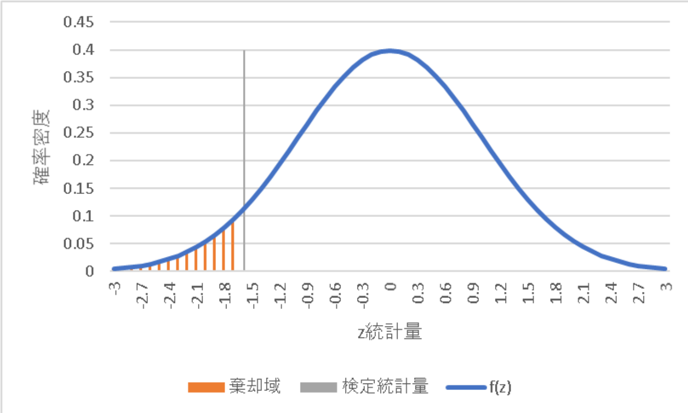
ニキビの数の不偏分散=0.4

標本の大きさ=5

z_α統計量の臨界値=1.64

ニキビの数の下限=8.95= $\mu - z_\alpha \frac{\sigma}{\sqrt{n}}$

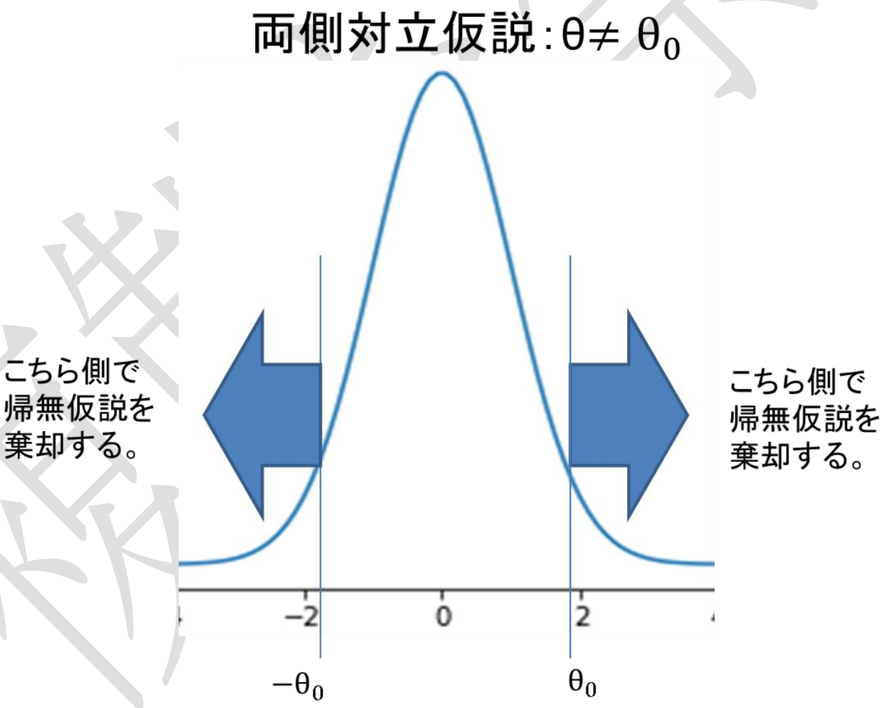
したがって、標本平均は 9 個なので t 統計量の下限よりも大きいので帰無仮説をすてることはできません。

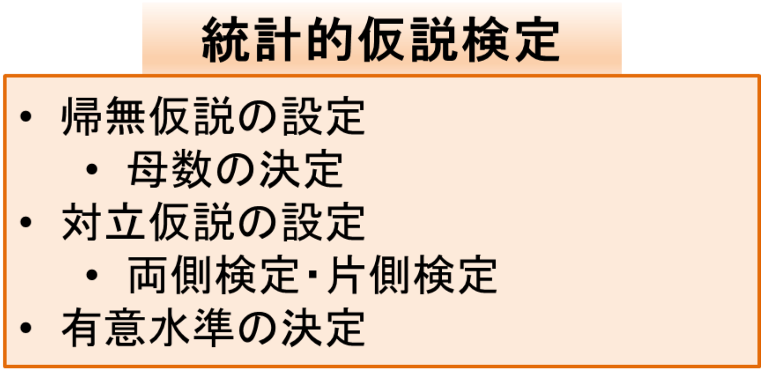


- コインのひずみ

　コインのひずみの有無を判断するために、表の出る確率を p=1/2 として、それを帰無仮説としました。この場合にひずみがあれば $p \neq 1/2$ となるので、対立仮説は $p \neq 1/2$ です。この際に、p は 2 分の 1 よりも大きい場合と小さい場合が可能です。したがって、このような対立仮説を両側対立仮説といいます。

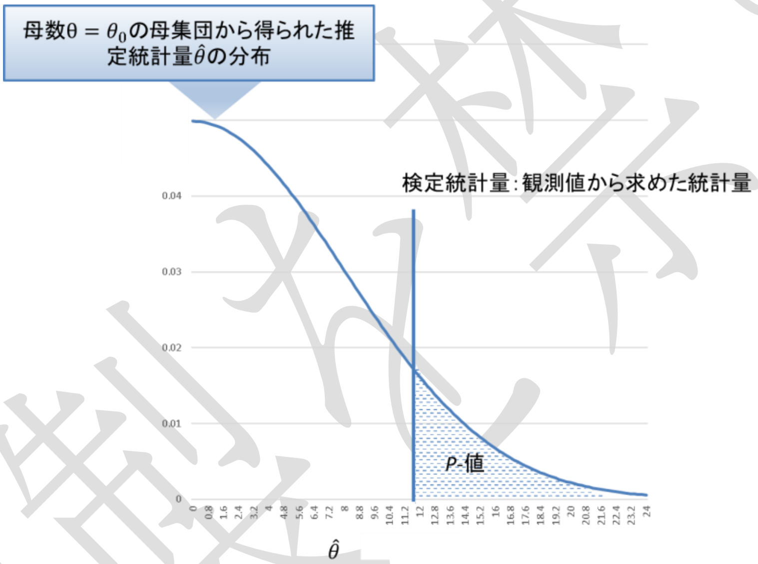
$$H_1: p \neq 1/2$$





5.2.6　p 値(有意確率)

p 値は、観察されたデータより極端な事象が現れる確率を表します。実現値の平均を $\bar{x}'$ とすると、p 値は $P(\bar{x} \geq \bar{x}')$ と書けます。観測されたデータをもとに棄却される有意水準を明確にできるので、固定した有意水準と観測データにもとづいた検定統計量を示すよりも、情報量が豊富であると考えられます。



一般に、

関係 (p は p 値を表す)	解釈
$0.01 \geq p$	帰無仮説を棄却する。
$0.1 \geq p \geq 0.01$	帰無仮説を棄却するに足る。
$p \geq 0.1$	帰無仮説を棄却するのは難しい。

「帰無仮説を棄却する」とは 0.01 以下の確率でしか起こらないことが起こった、ということです。
「帰無仮説の棄却は難しい」は棄却するに十分な証拠がないということです。
統計学の目的は極力誤った判断を減らすことにあります。

5.2.7 第一種の誤り

第一種の誤りとは、帰無仮説が正しいときに棄却してしまう誤りことです。この確率を α で表します。

5.2.8 第二種の誤り

対立仮説が正しいときに帰無仮説を受容してしまう誤りのことです。この確率を β で表します。対立仮説が正しい時に正しい確率は 1-β になります。

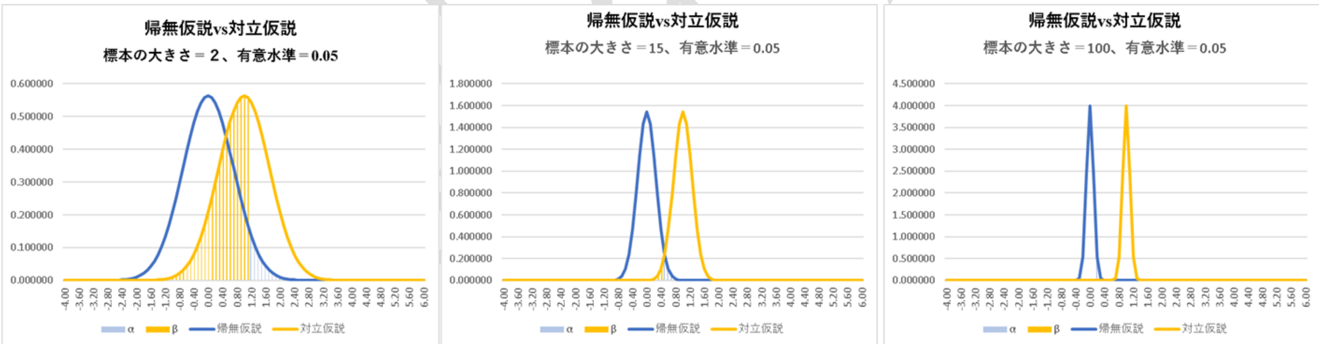
5.2.9 有意水準

第一種の誤りと第二種の誤りの両方を同時に小さくする必要があります。ところがこの二つは二律背反の関係にあります。そこで一般的には、第一種の誤りを一定以下の α に抑えながら、第二種の誤りを小さくするよう試みられます。この α は有意水準と呼ばれ、第一種過誤の確率を表します。

また、α を小さくすると β が大きくなるのですから、1-β は小さくなってしまいます。

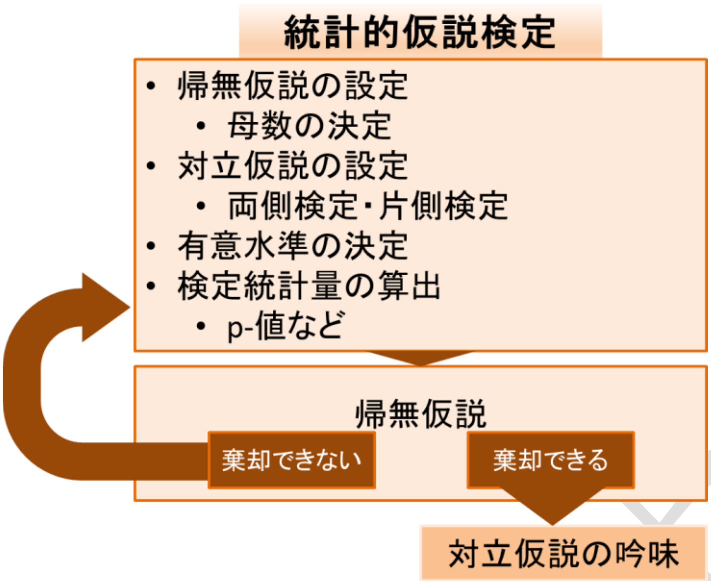
	H_0 が正しい (H_1 が誤り)	H_1 が正しい (H_1 が誤り)
H_0 の棄却	第一種の過誤 (α)	正しい判断 ($1-\beta$)
H_0 の採択	正しい判断 ($1-\alpha$)	第二種の過誤 (β)

例題 5.3： 帰無仮説の母平均がゼロ、母分散が既知で１、標本の大きさを 2, 15, 100、有意水準 0.05 として、標本が母集団から得られた場合の第一・二種の過誤の領域を標本平均の分布図の上で示してみましょう。対立仮説は母平均を１，母分散は既知で１とします。



標本の大きさが大きくなると２つの分布の重なる領域が小さくなるのが分かります。これは標本平均の標準誤差の影響です。

- 母平均が大きくなると、２つの分布は明確に異なるものとなります。
- 不偏分散が小さくなると、分母の標準誤差が小さくなります。
- 標本の大きさが大きくなると、標準誤差が小さくなります。



練習問題 5.1　帰無仮説を立てるときには$\boldsymbol{\mu} = \boldsymbol{\mu}_0$というように＝を使って仮説を立てます。$\boldsymbol{\mu} > \boldsymbol{\mu}_0$とか$\boldsymbol{\mu} < \boldsymbol{\mu}_0$とか$\boldsymbol{\mu} \neq \boldsymbol{\mu}_0$と立てないのはなぜでしょうか？