第 4 章　統計的推定

確率変数と確率分布を基礎とする推測統計を学習します。ここでは、未知の母数を、観測値(得られたデータ)をもとにもとめていきます。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のだろうかと考える必要があるのです。母集団から手元にあるデータ $X_1, X_2, \ldots, X_n$ 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、かつその値は変化する可能性があります。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 $X_1, X_2, \ldots, X_n$ から得られる 1 つの値で推定したい母数を示す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 ˆ 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。

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、

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

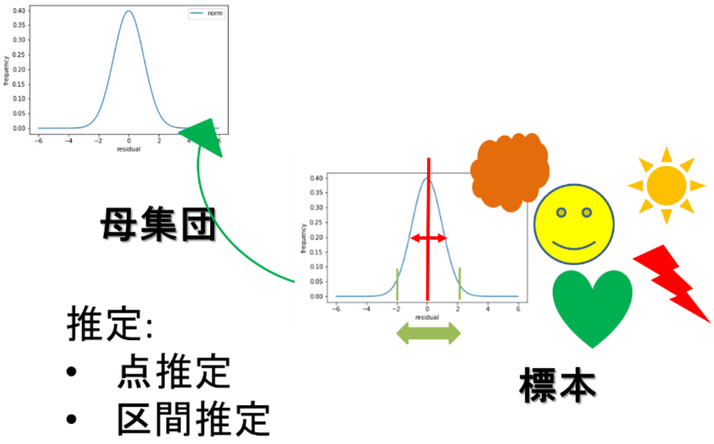
と表現します。

－ 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－α 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

平均値の区間推定を行ってみましょう。標本 $X_1, X_2, \ldots, X_n$ 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

　－ 正規分布を用います。

　－ z_α： 確率 α における標準正規分布の臨界値

α は有意水準で、1 － α が信頼係数となります。

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

－ 母分散が未知の場合

　－ t 分布を用います。

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

$$\bar{X} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$$

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、

$$z=\frac{\mathrm{var}(X)\cdot(n-1)}{\sigma^2}\sim\chi^2_{n-1}$$

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

$$\frac{\sum(X_i-\bar{X})^2}{\chi_{\alpha/2}(n-1)}<\sigma^2<\frac{\sum(X_i-\bar{X})^2}{\chi_{1-\alpha/2}(n-1)}$$

となります。標本の大きさが大きくなると信頼区間は狭くなります。

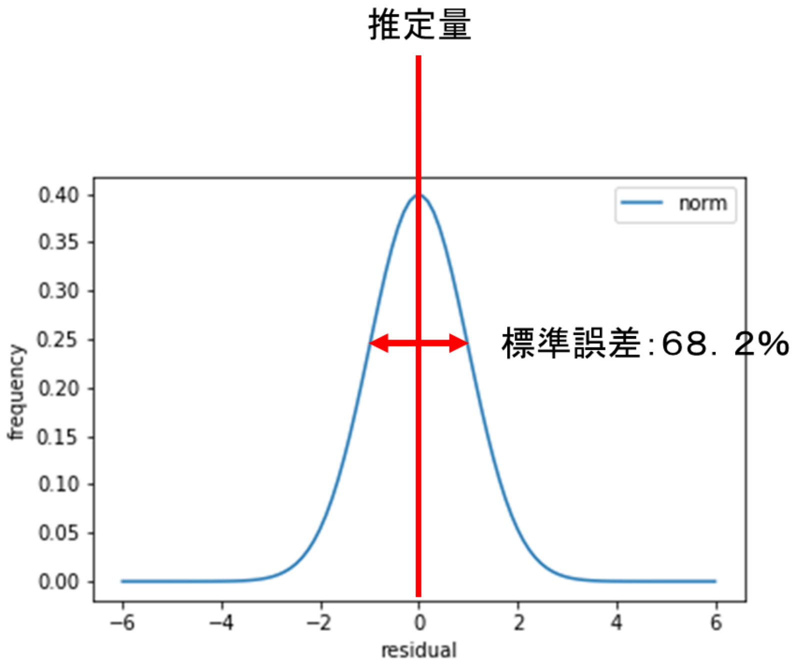
例：エクセルによるワインデータの主な要素の区間推定

	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
4	最大値	15.90	1.58	1.00	15.50	0.61	72.00	289.00	1.00	4.01	2.00	14.90	8.00
5	最小値	4.60	0.12	0.00	0.90	0.01	1.00	6.00	0.99	2.74	0.33	8.40	3.00
6	第1四分位範囲	7.10	0.39	0.09	1.90	0.07	7.00	22.00	1.00	3.21	0.55	9.50	5.00
7	第3四分位範囲	9.20	0.64	0.42	2.60	0.09	21.00	62.00	1.00	3.40	0.73	11.10	6.00
8	範囲	11.30	1.46	1.00	14.60	0.60	71.00	283.00	0.01	1.27	1.67	6.50	5.00
9	歪度	0.98	0.67	0.32	4.53	5.66	1.25	1.51	0.07	0.20	2.42	0.86	0.22
10	尖度	1.13	1.22	-0.79	28.55	41.57	2.03	3.79	0.93	0.81	11.68	0.20	0.29
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		酒石酸濃度	酢酸濃度	クエン酸濃度	残糖濃度	塩化ナトリウム濃度	遊離二酸化硫黄濃度	総二酸化硫黄濃度	密度	pH	硫化カリウム濃度	アルコール度数	評価
16		7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	5

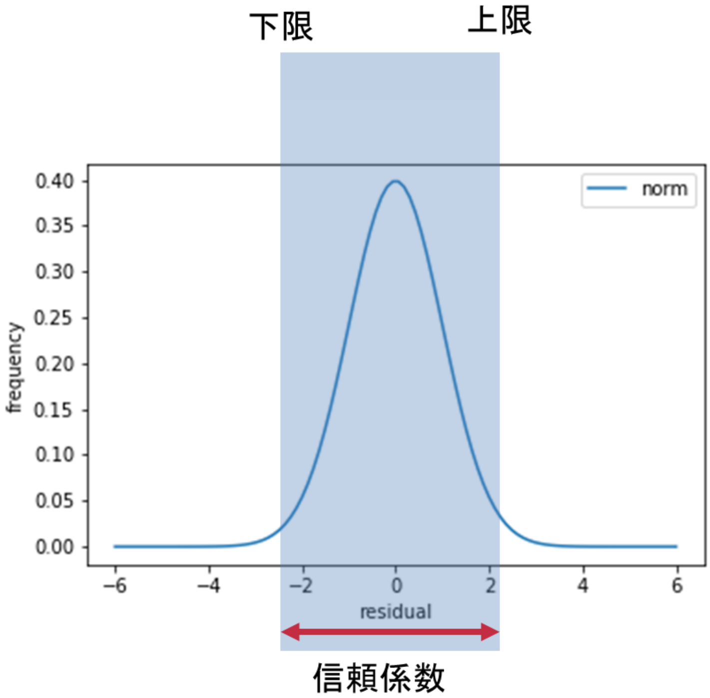
信頼区間についてのエクセル関数

	A	B
13	自由度	=+COUNT(B18:B1616)-1
14	信頼係数	0.95
15	上限	=B3+B5*T.INV(B14,B13)
16	下限	=+B3+B5*T.INV(1-B14,B13)
17		酒石酸濃度
18		7.4

- 点推定
  - 推定量
  - 標準誤差



- 区間推定
  - 上限、下限
  - 信頼係数



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よ。

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。

何度も何度も繰り返し、データを得て、まず仮説をたてます。しかし、得られる値が、いつでもその仮説からは想定出来ないのであれば、実際のデータは想定する仮説とは違うかも知れません。推測統計の手法の 2 つ目は統計的仮説検定です。これは、与えられた仮説に対して得られたデータからその正しさを統計的に判断する方法です。

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ないし、確率としてあり得るのかもしれません。判断が難しいときがあります。

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、確率的に起き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き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。そしてこのような判断のしかたが**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。

5.2 仮説検定の構造

帰無仮説と対立仮説を立て、検定に用いる統計量を定めて、標本をもとに帰無仮説の有無を判断します。

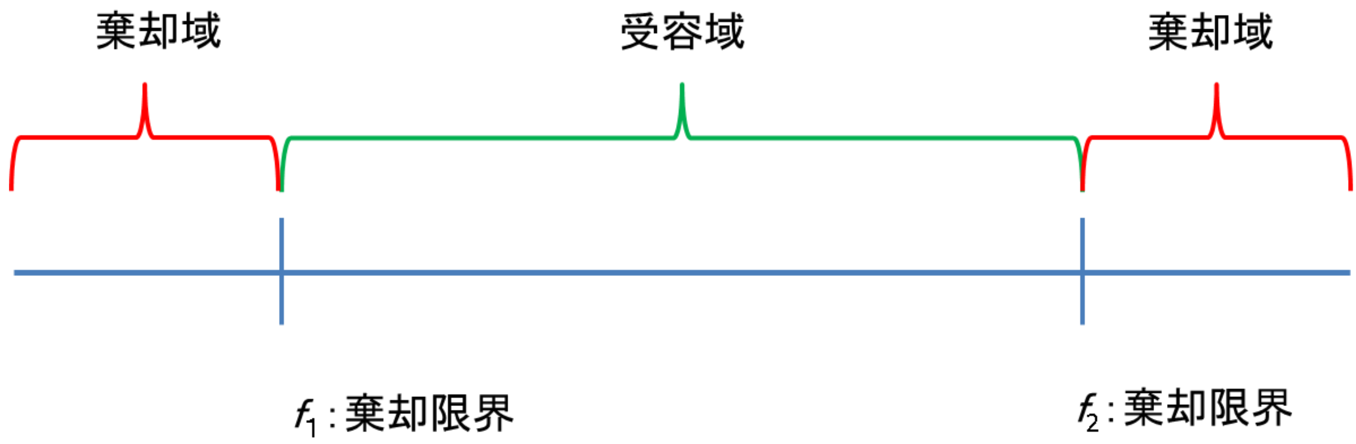
5.2.1 帰無仮説と対立仮説

「効果がある」とか「差がある」という検定をしたいときに、仮説検定では、「差がない」とか「効果がない」とかという仮説をたてます。それを帰無仮説といい、H_0 と書きます。それに対する仮説を対立仮説といいます。それを H_1 と書きます。実際には H_0 は棄却されてほしいのです。

- 対立仮説には片側検定と両側検定があります。

統計的仮説検定において確率変数 X_i が得られたとします。確率変数 X が取り得る全事象、または標本空間は、X_i の関数として表現されるある統計量 $F(X_i)$ によって 2 つの領域に分割されます。この統計量を検定統計量といい、確率変数は 2 つの領域に分けられます。1 つは受容域、もう一方は棄却域です。棄却域とは、帰無仮説が棄却される統計量がとる領域のことです。

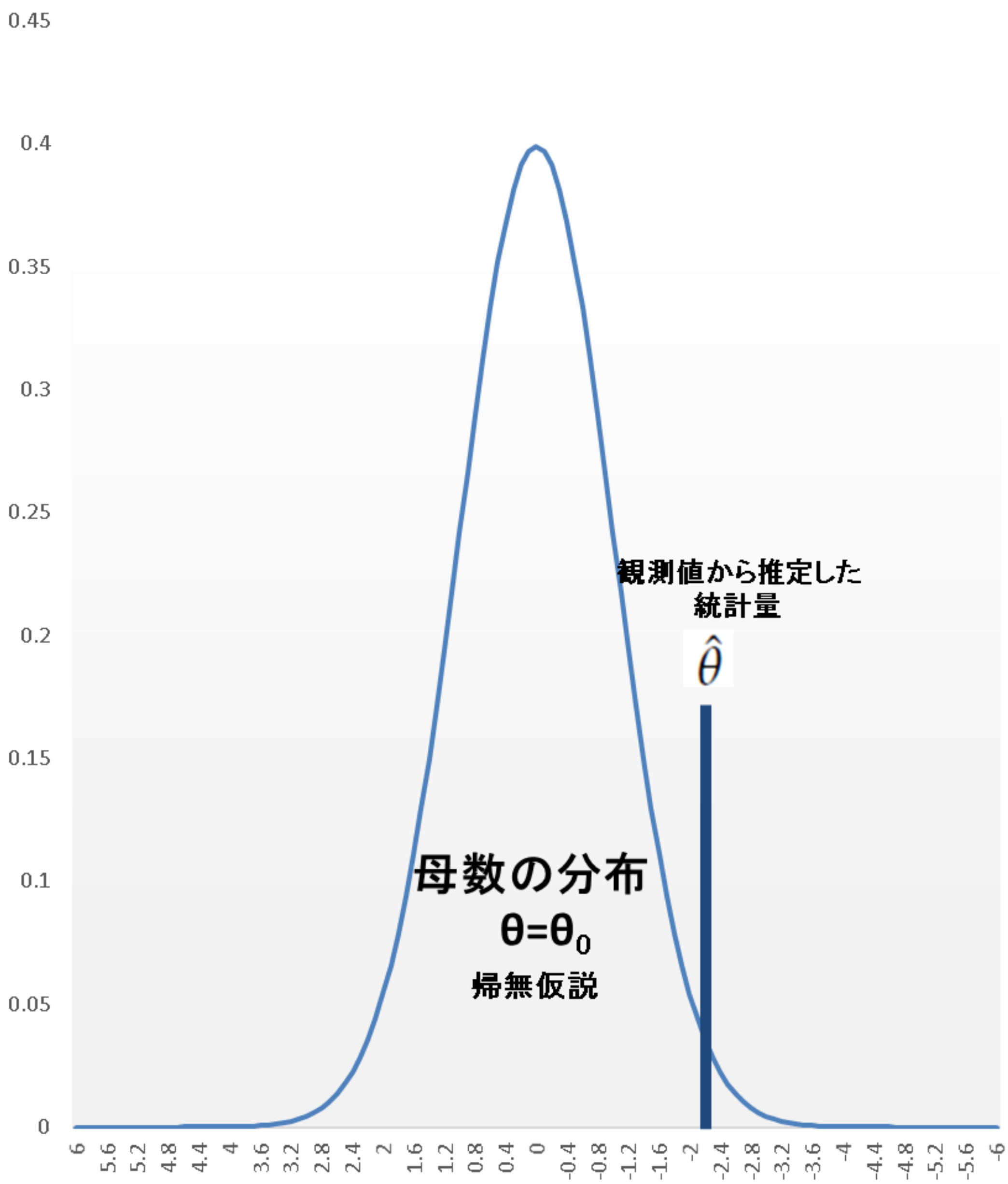
－ 棄却域は片側検定の場合、$\{X_i | F(Xi) < f_1\}$、または $\{X_i | F(X_i) > f_1\}$、両側検定の場合、$\{X_i | F_i\ (X) < f_1\}$ または $\{X_i | F(X_i) > f_2\}$ などと表されます。f_1，f_2 を棄却限界といいます。



5.2.2 帰無仮説

確率変数 X_i が得られたとき、この X_i はもととなる集団、つまり母集団から得られたと仮定します。この母集団はある確率分布にしたがっている必要があります。この確率分布は統計量 θ により規定されます。それを θ_0 とします。したがって、X_i がこの母集団から得られたかどうかを特定するために統計量を用いることができます。その得られた値について帰無仮説 (H_0) を立てます。

手元に得られたデータが x_i のとき, その標本から母数を推定できます。この推定量が得られる確率を母数の確率分布から推定します。この確率が有意水準 α より小さければ、H_0 に対してかなりまれな事象が起きたとして、帰無仮説を棄却します。そうでなければ棄却はしません。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。



　最も簡単な例は得られたデータの母集団の平均がゼロであることが既知の場合です。その場合の帰無仮説は H_0: $\mu = 0$ です。手元にあるデータから得られた平均の推定量が有意水準よりも小さく稀な事象であれば、帰無仮説は棄却されます。

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。

$$H_0: p = 0.5$$

- 薬の効果

あるニキビの薬の効果を確かめるために、無作為に選んだ患者に薬を処方することにします。選ばれた患者のニキビの平均個数は 10 とします。薬に効果があれば薬の投与後にニキビの数は減るはずです。効果が無ければニキビの数に変化はありません。そうすると帰無仮説は患者のニキビの数の平均 μ=10 となります。

$$H_0: \mu= 10$$

- 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しょう。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 帰無仮説と母数

これらの例の表の出る確率 p=1/2、患者の平均ニキビ数 μ=10、成功確率 p=1/2 はどれも母数と呼ばれます。得られた確率変数 Xi がある分布にしたがうとは、これらの母数により規定されている分布にしたがうということです。このとき、この母数が属する空間を母数空間と呼び Θ で表します。確率変数 X が母数 θ によって規定される分布にしたがうという仮説をたてるとき、この仮説を帰無仮説といいます。

帰無仮説が正しくないと結論付けるとき、帰無仮説を棄却するといいます。仮説が正しくないという判断には至らないとき、帰無仮説は「棄却するには十分ではない」とします。

5.2.3 対立仮説

帰無仮説が棄却されたときに採用される仮説が対立仮説です。

検定には２つのタイプがあります。1 つは、差の大小を比べたいときです。購入したミルクの量が表示よりも少なそうだと思ったときには、少ないほうに大きくずれていることが判断の対象になります。そのような検定を片側検定といいます。

もう 1 つは明確に差があるか無いかを検定したいときです。その場合に、差の有無だけを判断したいのですから、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。

片側対立仮説と母数

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

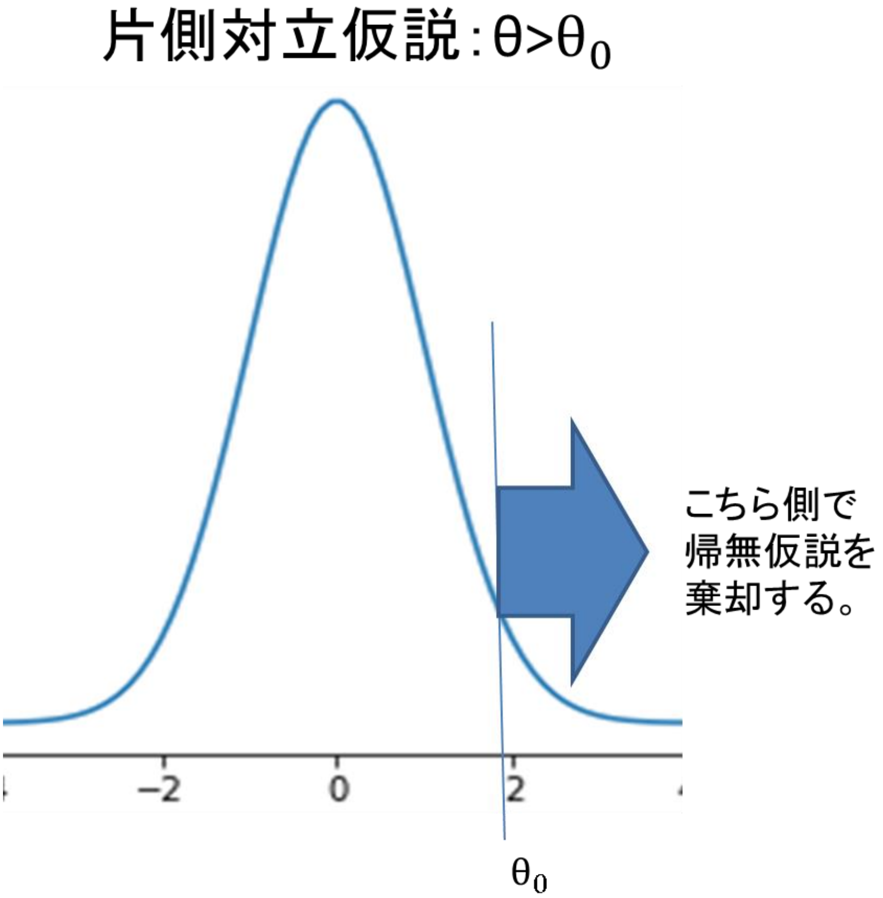
両側対立仮説と母数

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

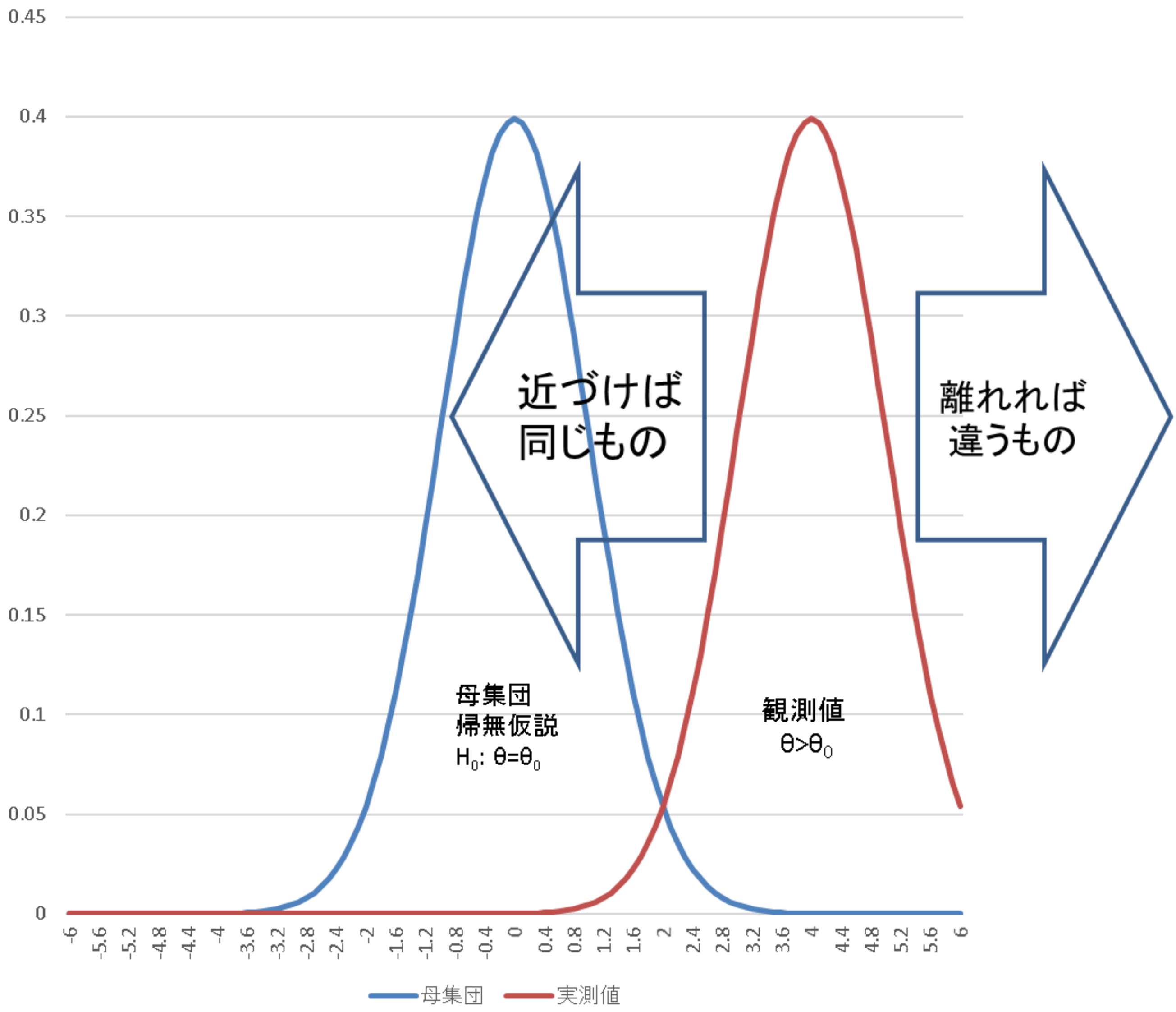
帰無仮説は比べたいものを等しいとして、対立仮説で片側検定、両側検定ので検定のタイプを指定します。例を見てみましょう。

- 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。

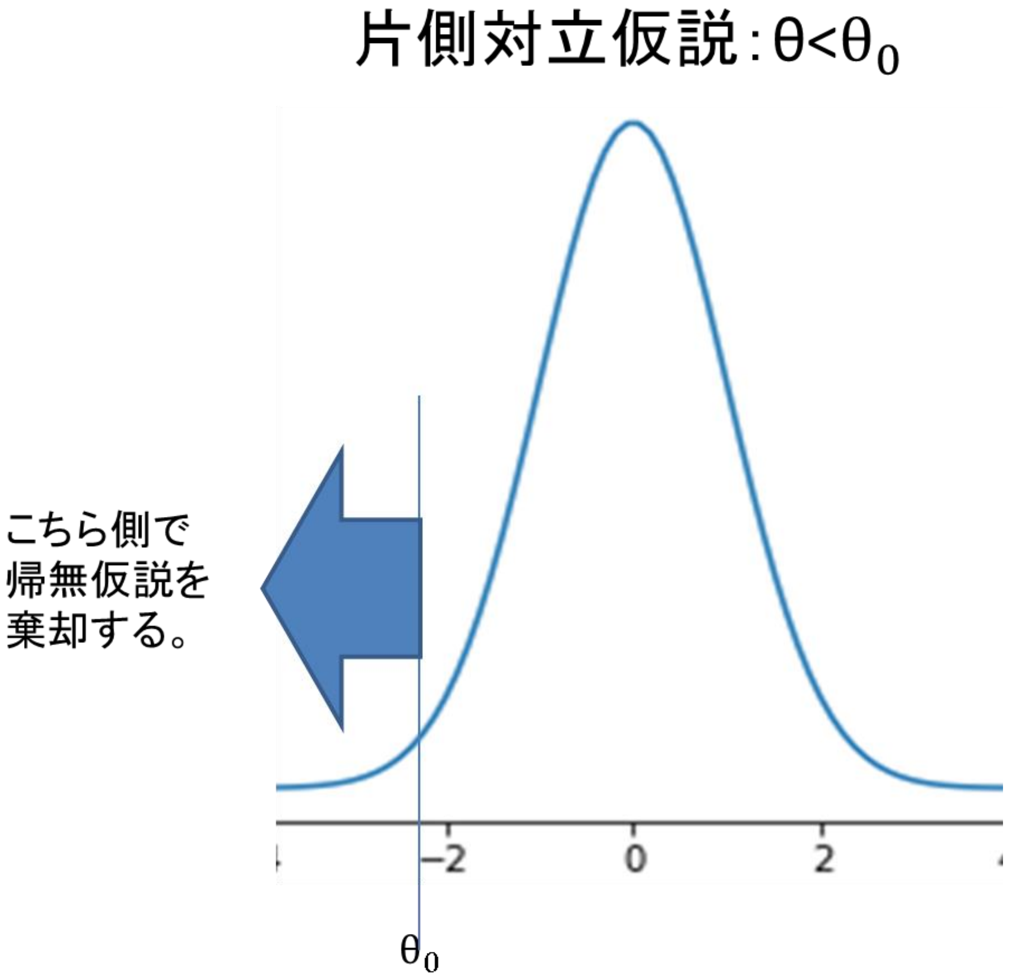


H_1: p>1/2



- 薬の効果

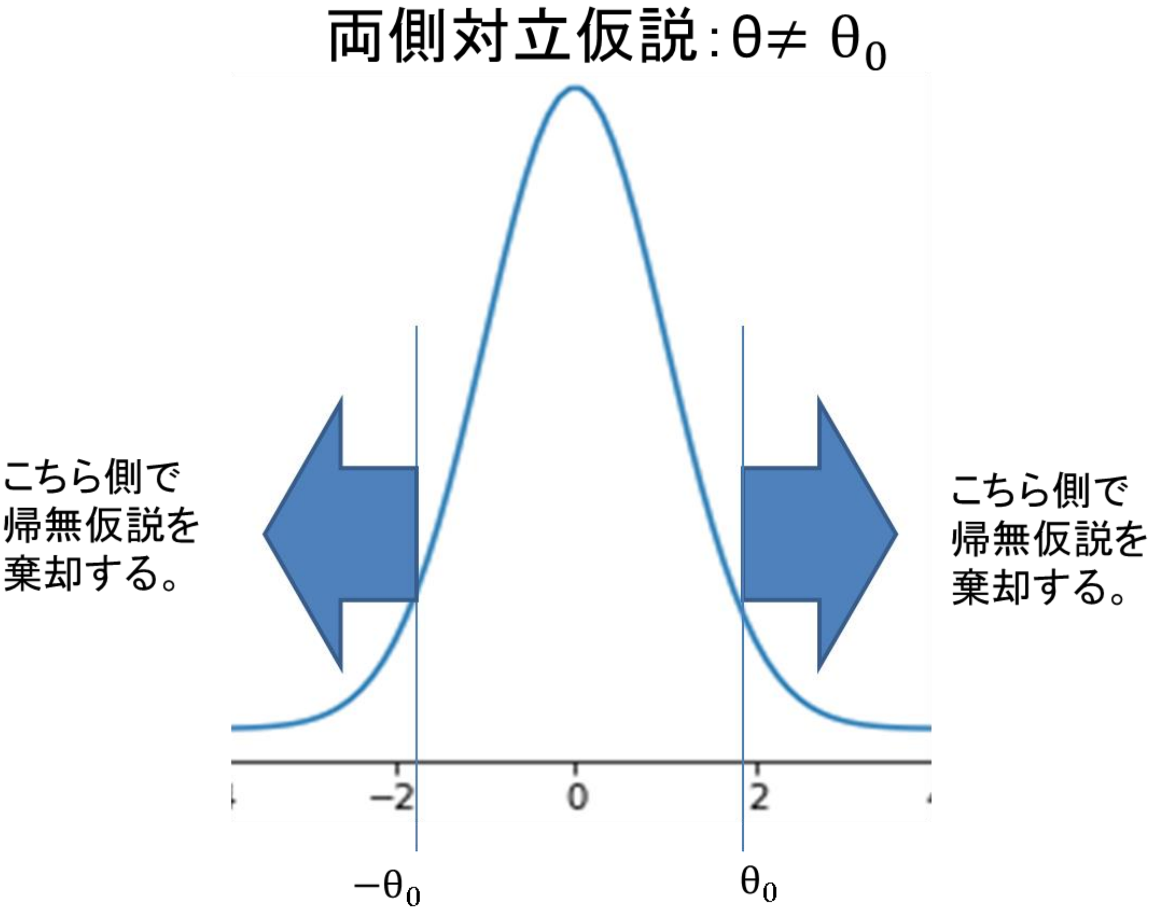
あるニキビの薬の効果を確かめために、帰無仮説は患者のニキビの数の平均 $\mu=10$ としました。薬の投与の後に平均のニキビの数が増えたのでは薬の効果は逆効果です。平均のニキビの数は減らなくてはなりません。したがって、対立仮説は $\mu<10$ となり、分布の裾の片側だけを考えればよいことになります。これが片側対立仮説です。



H_1: μ<10

- 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≠1/2 となるので、対立仮説は p ≠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: $p \neq 1/2$

5.2.4 帰無仮説が棄却されるという意味

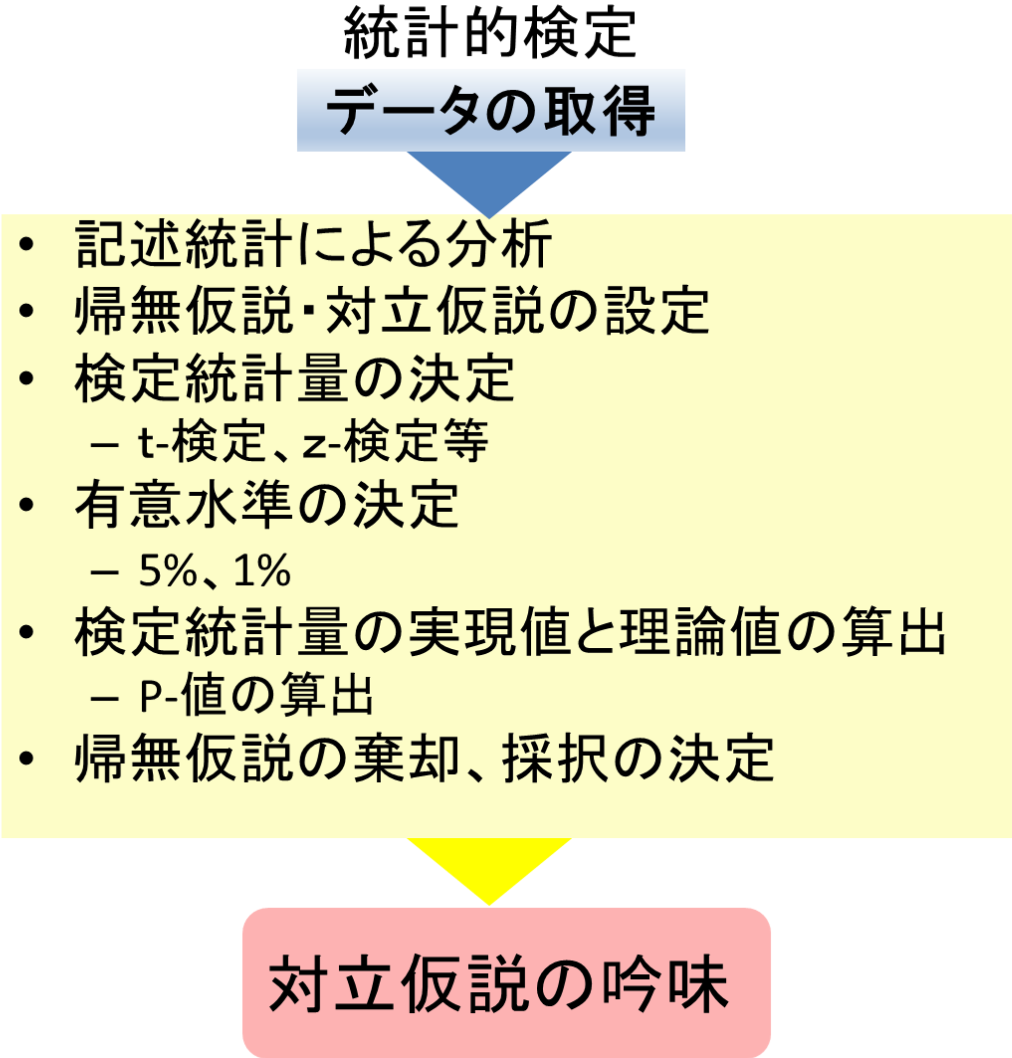
仮説検定では標本にもとづいて帰無仮説と対立仮説のどちらかを統計的な判断で選択しています。対立仮説が選択されたというときには、この選択が誤りであるという確率は α 以下であると保証されています。つまり、対立仮説が強く成り立っていると主張することができます。

もちろん、確率変数が、帰無仮説が仮定する条件を満たしていないために棄却された場合には、この限りではありません。

5.2.5 帰無仮説が棄却されないの意味

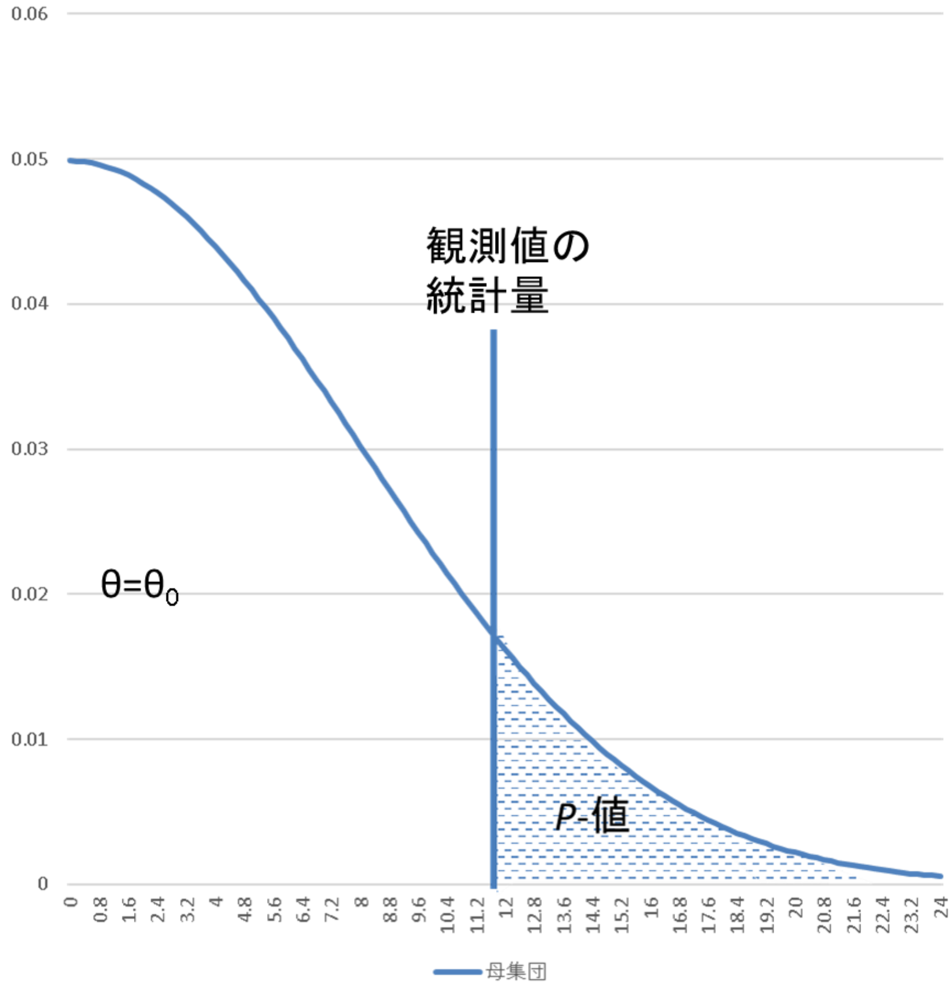
帰無仮説が棄却されなかったからといって、それを積極的に支持する理由にはなりません。それが誤りである確率が低いということはどこにも言及されていないのです。つまり、帰無仮説が棄却されないからといって、強く支持することはできないのです。帰無仮説が単に棄却される理由が不十分であったに過ぎないのです。ですので、帰無仮説を受容するとは言わずに、棄却するには十分ではないと表現するのです。

ここで注意が必要なのは確率変数が独立でないとか、一様でないとかという理由で棄却されてしまう場合があります。確率変数 X がある確率分布にしたがうとしたときは、確率変数は条件を満たしていなければなりません。確率変数 X がそれを満たさなければ、X はその確率分布にしたがいません。そうすると統計的検定には意味がなくなります。



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。

両側検定の場合の p 値の使用には注意が必要です。片側検定の p 値の 2 倍とする方法と、観測データから算出した出現確率よりも、出現する確率が小さい事象の確率とする方式があります。後者の場合には、左右対称でない分布では前者と異なる結果となります。

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

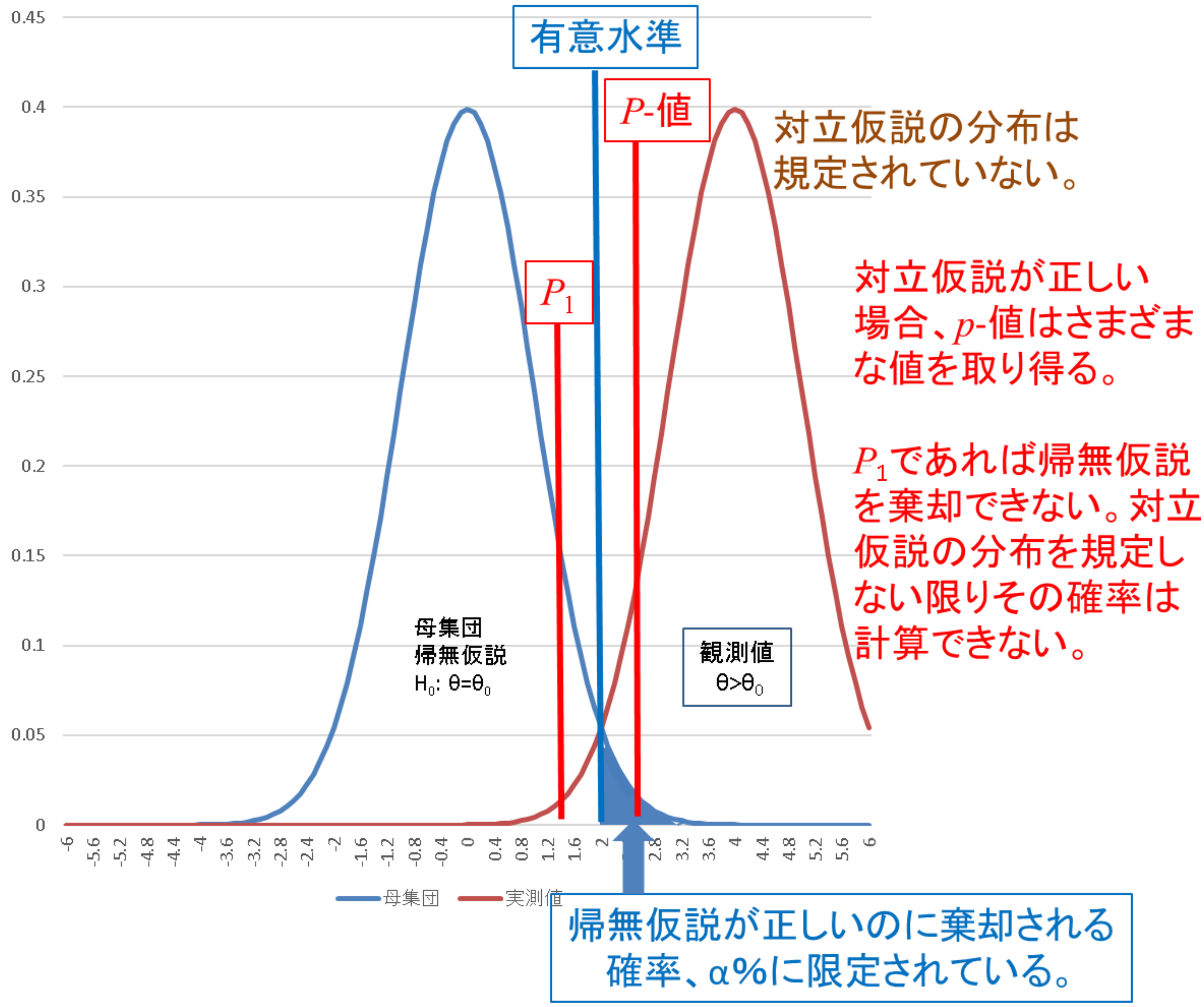
対立仮説が正しいときに帰無仮説を受容してしまう誤りのことです。この確率を β で表します。対立仮説が正しい時に正しい確率は $1-\beta$ になります。β は帰無仮説、対立仮説のかなでは規定されていません。++

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。しかし、これは帰無仮説が正しくない確率ではないことに注意が必要です。β は帰無仮説からも対立仮説からも得られるものではありません。別途設定する必要があります。

また、α を小さくすると β が大きくなるのですから、$1-\beta$ 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)


有意水準
P-値
P1
対立仮説の分布は
規定されていない。
対立仮説が正しい
場合、p-値はさまざま
な値を取り得る。
P1であれば帰無仮説
を棄却できない。対立
仮説の分布を規定し
ない限りその確率は
計算できない。
母集団
帰無仮説
H0: θ=θ0
観測値
θ>θ0
帰無仮説が正しいのに棄却される
確率、α%に限定されている。
母集団
実測値
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0