

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ことがあります。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。判断が難しいことがあります。

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か。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。

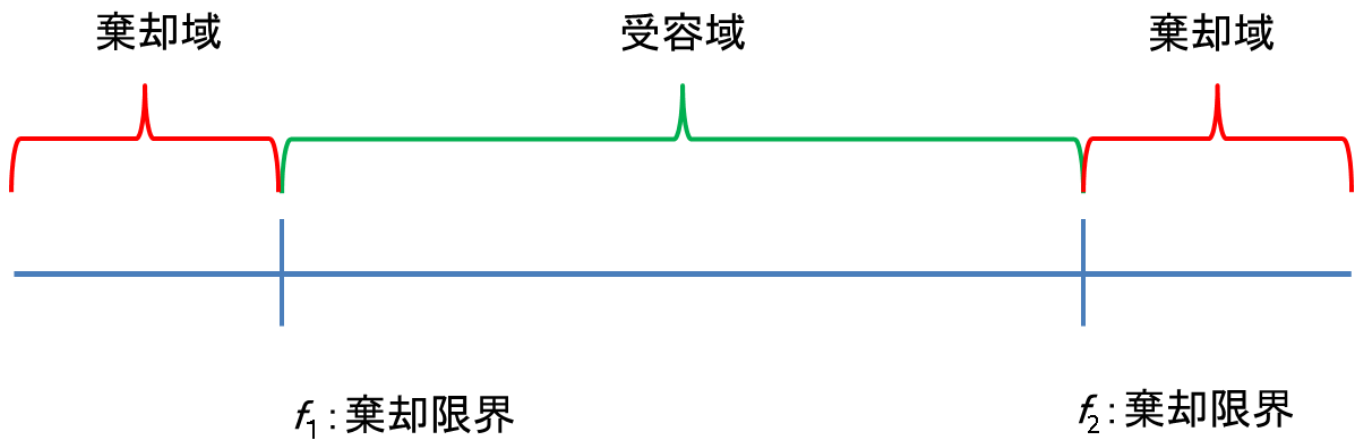
5.2.1 帰無仮説と対立仮説

「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。

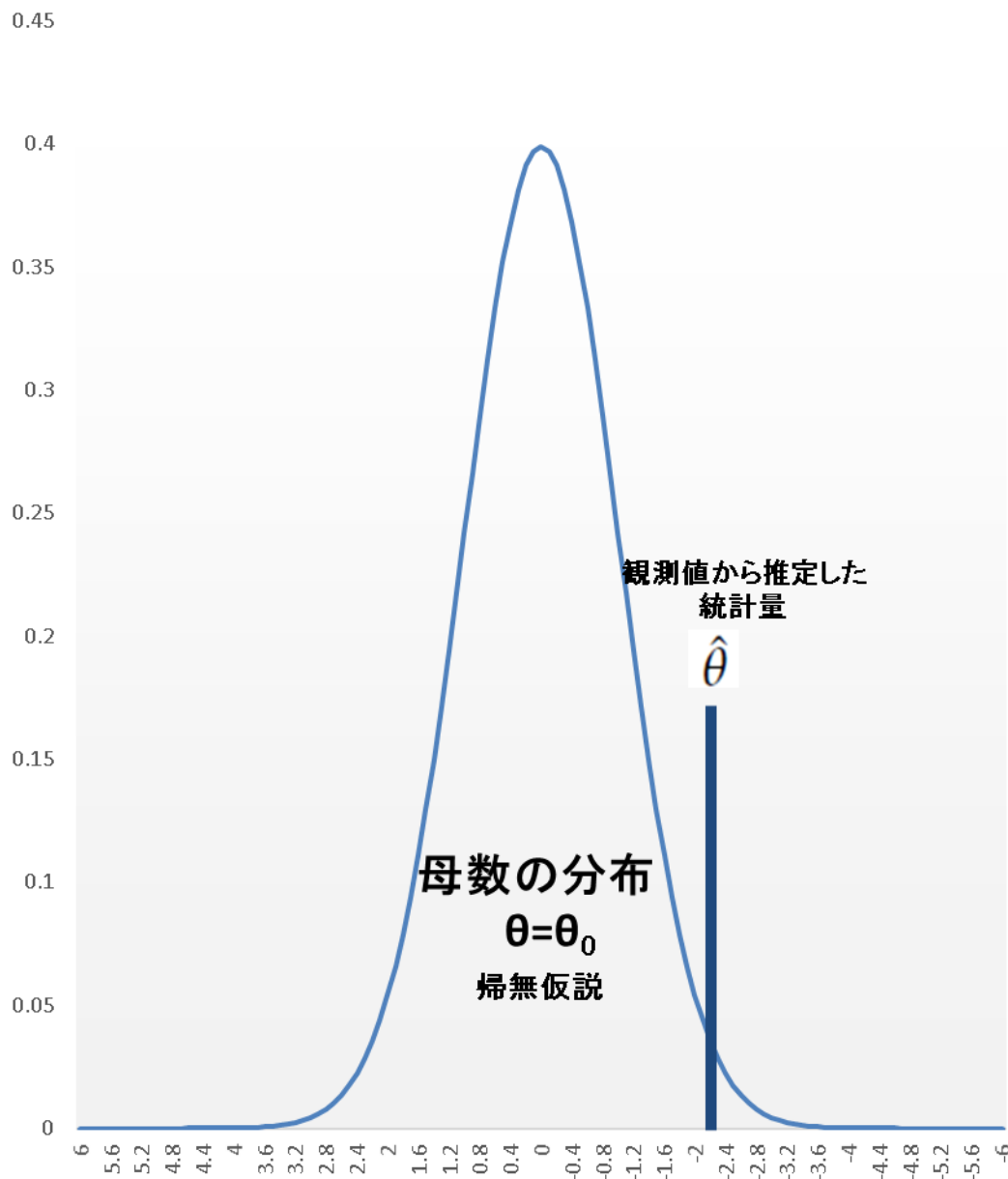
－ 棄却域は片側検定の場合、 $\{X_i | F(X_i) < f_1\}$ 、または $\{X_i | F(X_i) > f_1\}$ 、両側検定の場合、 $\{X_i | F_i(X) < f_1\}$ または $\{X_i | F(X_i) > f_2\}$ などと表されます。 f_1, f_2 を棄却限界といいます。



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われています。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が。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が、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 $\mu=10$ 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 $\mu=10$ 、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 X_i 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

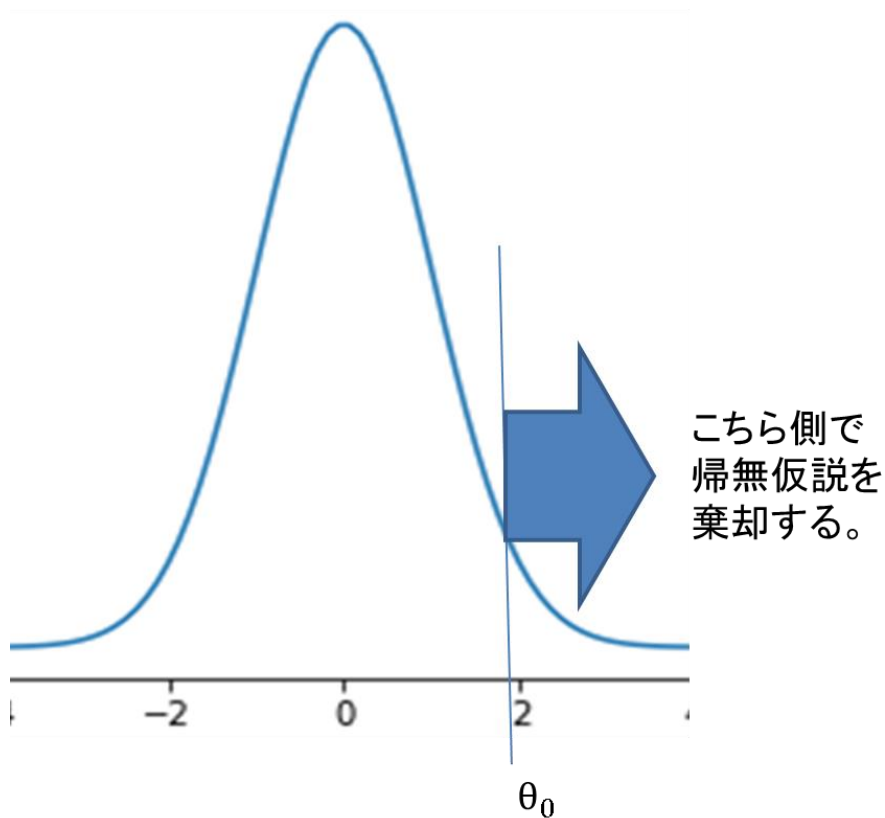
$$H_1: \theta \neq \theta_0$$

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。

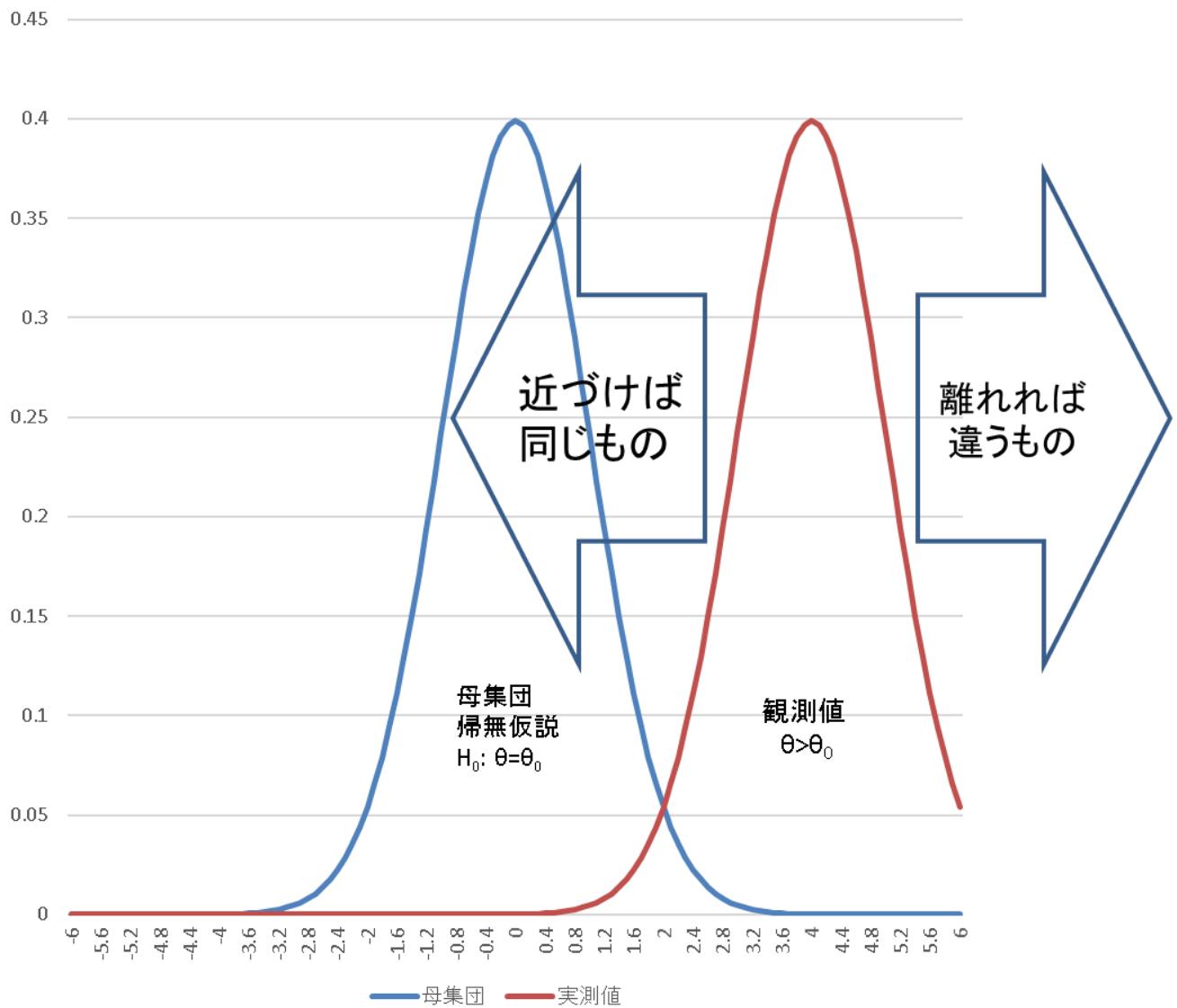
- クラゲの占い

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预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预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。

片側対立仮説: $\theta > \theta_0$



$$H_1: p > 1/2$$

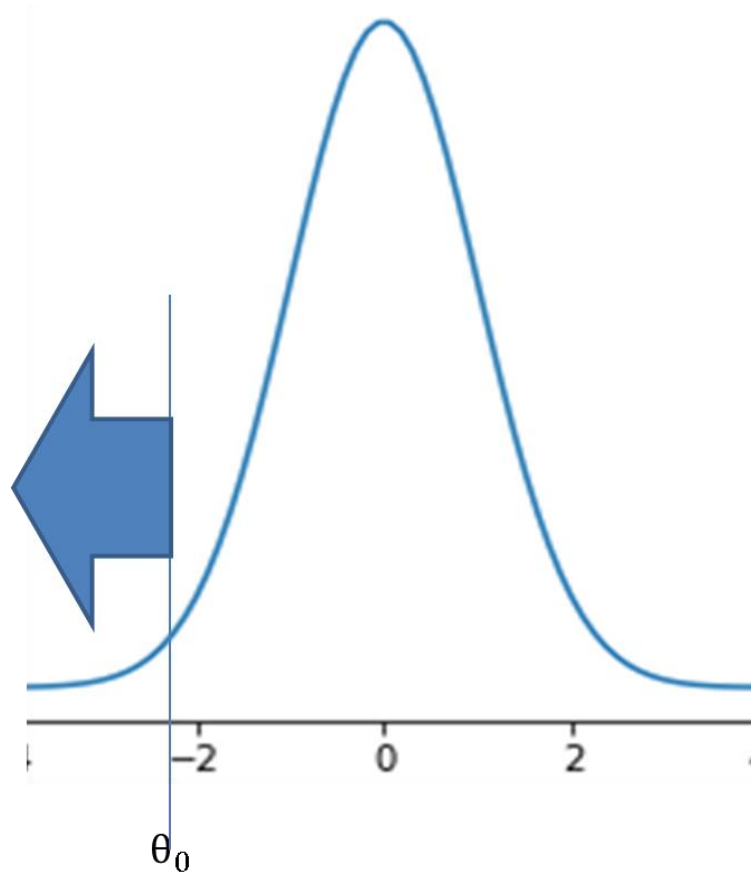


- 薬の効果

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。

片側対立仮説: $\theta < \theta_0$

こちら側で
帰無仮説を
棄却する。

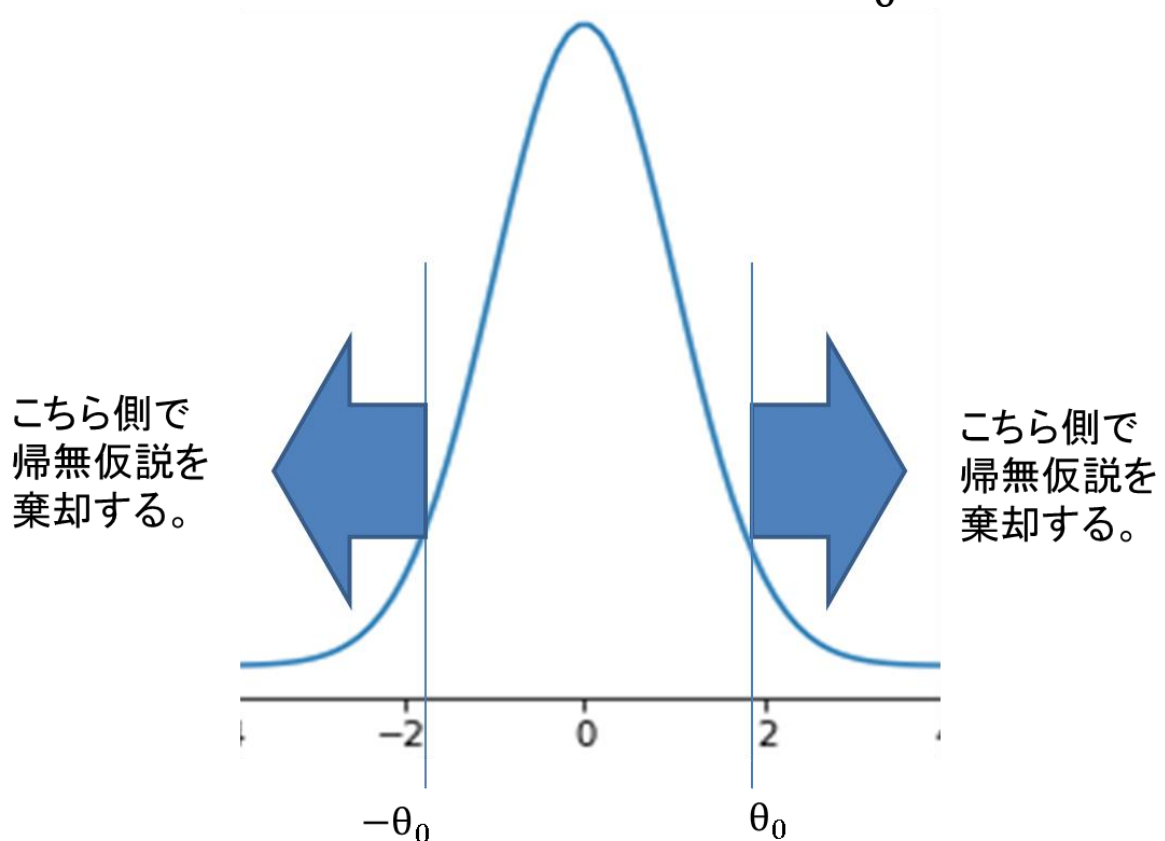


$$H_1: \mu < 10$$

- 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、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。

両側対立仮説: $\theta \neq \theta_0$



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

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ったときは、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。

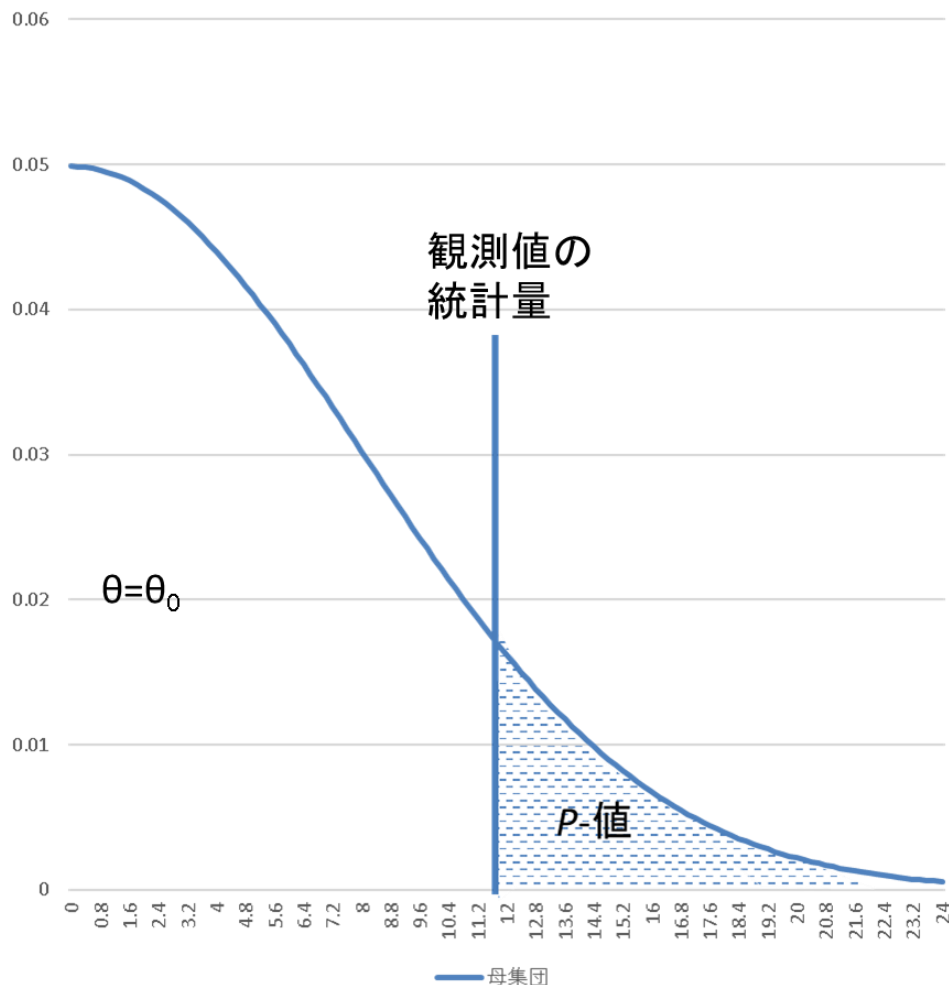
統計的検定 データの取得

- 記述統計による分析
- 帰無仮説・対立仮説の設定
- 検定統計量の決定
 - t-検定、z-検定等
- 有意水準の決定
 - 5%、1%
- 検定統計量の実現値と理論値の算出
 - P-値の算出
- 帰無仮説の棄却、採択の決定

対立仮説の吟味

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に試みられます。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)

