統計的な推測

正規分布⇒仮説検定

最終問題　君たちは今１億円を賭けた大勝負のまっただ中にいる。勝負の方法はいたって簡単。

【ルール】１枚のコインを２０回投げて、表が多く出た方の勝ち。イカサマが発覚した時点で負け。

【状　況】相手が先攻で、１５回表が出てしまった。なかなか難しい状況になった…。

君たちの選択肢は以下の２つ。

選択肢①：このまま１６回以上表が出ることを信じて戦う　⟹ 勝てる確率はいかほどか

選択肢②：相手のコインの偏りに言及してイカサマを暴きにかかる　⟹ 表裏に偏りアリと言えるか

※確率５％未満は非現実的であると考えてよい。

選択肢 $\boxed{}$ を選択

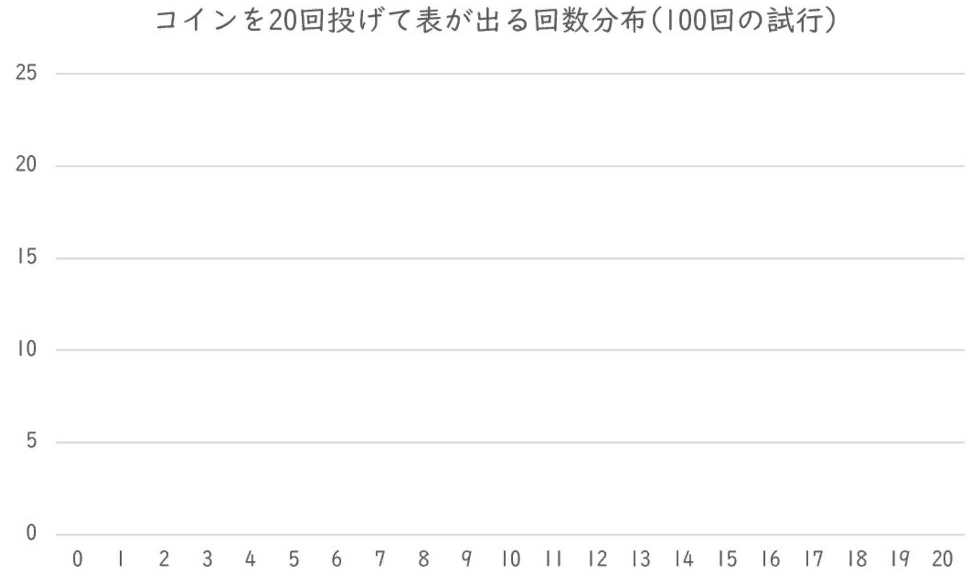
思考

結論

問題点・課題点　◎評価ポイント

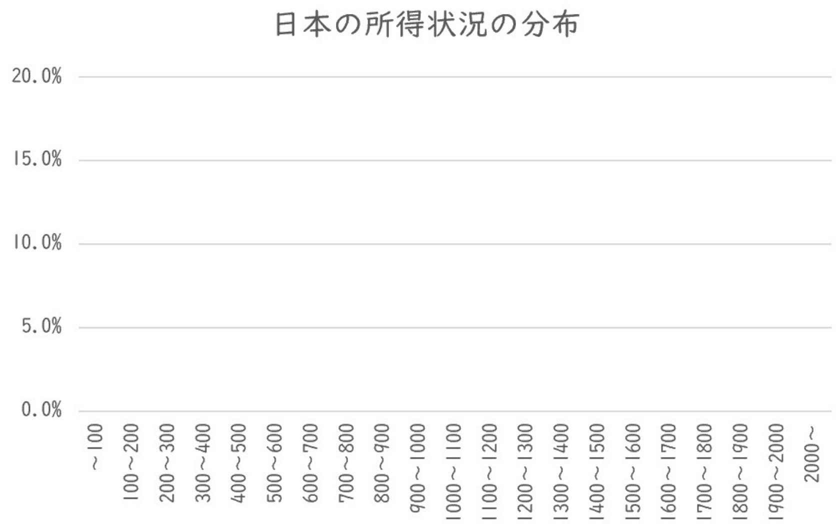
第１章　様々な分布　　※世の中にはさまざまな形をした分布が存在しています。少し触れてみましょう。

予想①



予想②

平均値：５３６万円
中央値：４１０万円



予想③

平均値：１７１ｃｍ
中央値：１７０ｃｍ

第２章　非常に美しい分布「正規分布」

このような釣鐘型（ベル型）の分布を「正規分布」と呼び、「正規分布 $N(m,\sigma^2)$ に従う」と言います。

なんと自然界の多くのデータがこの正規分布に従っていることが知られており、非常に実用的です。

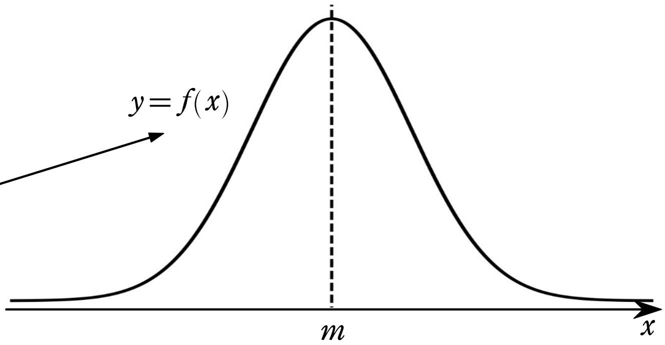
$Point.$ **（正規分布）**

正規分布 $N(m\,,\sigma^2)$

※ m；期待値(平均)，σ^2 ：分散

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$$

※この式は覚えなくてよし！

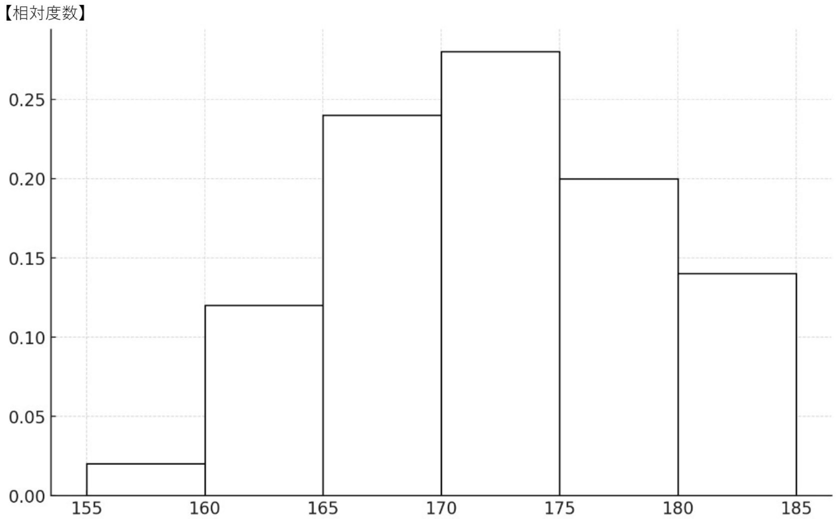


ではこの正規分布の何がすごいのかを理解するために、まずは基礎知識を学んでおきましょう。

第３章　確率は面積だ！！

ある１００人の身長を調べた結果、以下の度数分布表，ヒストグラムになった。

身長の階級(cm)	度数	相対度数
155～160	2	0.02
160～165	12	0.12
165～170	24	0.24
170～175	28	0.28
175～180	20	0.20
180～185	14	0.14
計	100	1.00



※これって要は $\boxed{}$ のこと！

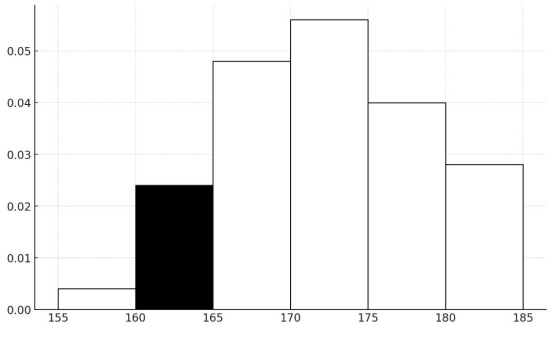
※当たり前のことですが、すべての相対度数を足すともちろん「１」になります。

そこでヒストグラム全体の面積を「１」とすると

↓

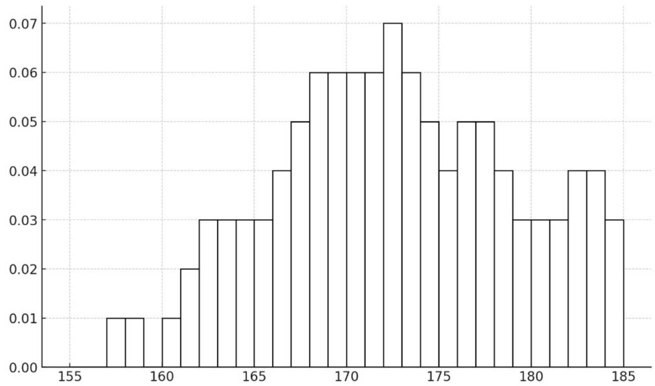
「160～165cm」の部分を表している部分の長方形の面積(黒塗部)が

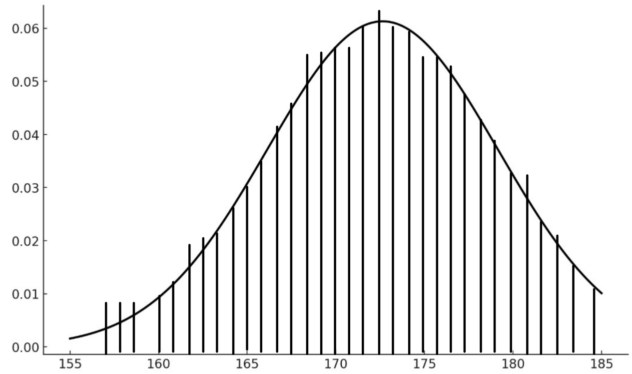
その階級が選ばれる確率、すなわち0.12を表していることになります。
（相対度数）



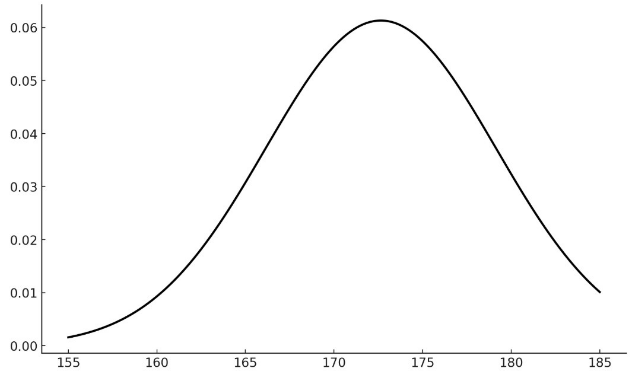
これを確率変数 X を用いて、$\boxed{P(160\leqq X\leqq 165)=0.12}$ と表現します。

ちなみにこのデータをどんどん細かくしていくと…





ヒストグラム削除！



なんと最後は曲線に！

この曲線　⇒「分布曲線」

この関数 $f(x)$　⇒「確率密度関数」

例題１． 上の問題において、以下の問いに答えよ。

(1) $170 \leqq X \leqq 180$ である確率を求める式を立式せよ。

(2) $\displaystyle\int_{155}^{185} f(x)\,dx$ の値を求めよ。

この章をまとめると…

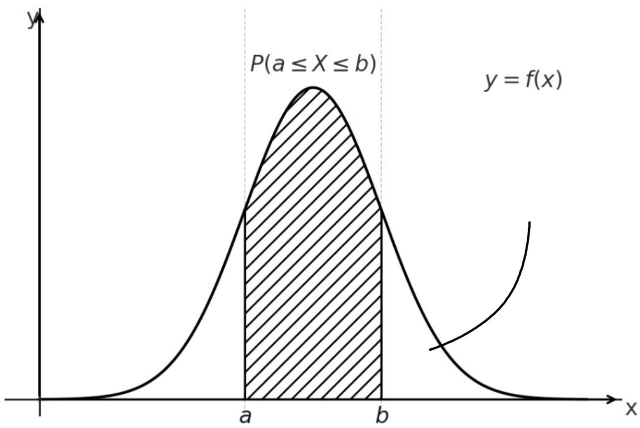
$Point.$ **（確率は面積だ！）**

① $P(a \leqq X \leqq b) = \displaystyle\int_a^b f(x)\,dx$

※確率は面積なので、もちろん積分利用！

② $\alpha \leqq X \leqq \beta$ のとき $\displaystyle\int_\alpha^\beta f(x)\,dx = 1$

※確率全部足すと「１」になるから当たり前！



☆この感覚を持つことで、正規分布を最大限利用できるようになります。

次は正規分布の活用です。さあ正規分布の凄さを探しにいきましょう。

第４章　「正規分布」を活用しよう！　※第３章の考えを忘れずに！

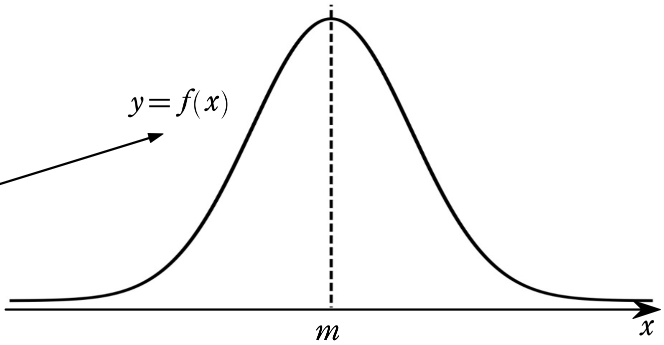
$Point.$ **（正規分布）**

正規分布 $N(m\,,\sigma^2)$

※ m ：期待値(平均)， σ^2 ：分散

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$$

※この式は覚えなくてよし！



問題１． 確率変数 Z が正規分布 $N(0,\ 1)$ に従うとき，次の確率を求めよ。

① $P(0\leqq Z\leqq 1.23)$

② $P(-2\leqq Z\leqq 1.23)$

③ $P(1\leqq Z\leqq 2)$

④ $P(2\leqq Z)$

問題２． 確率変数 X が正規分布 $N(170.5\,,\ 5.4^2)$ に従うとき，確率 $P(X\geqq 178)$ を求めよ。

解いてみた感想　◎評価ポイント

第５章　「標準正規分布」って最高！　※第３章の考えを忘れずに！

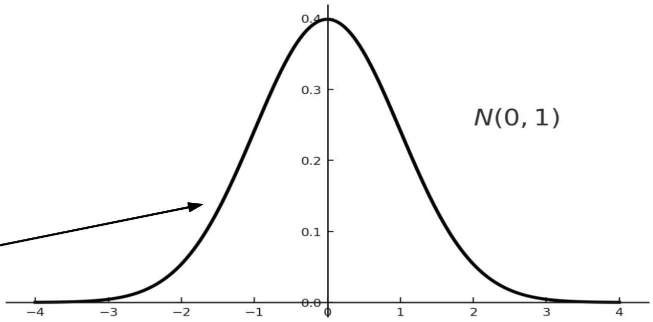
第４章は苦しかったですね。恐ろしく複雑な計算になりました。でも実は…正規分布 $N(0,\ 1)$ に関してのみ、面積を計算してくれた偉大な先人がいるのです。そのような特別な分布のことを次のように呼んでいます。

$Point.$ **（標準正規分布）**

正規分布 $N(0\,,1)$ のことを標準正規分布という

※ $m=0$，$\sigma=1$

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

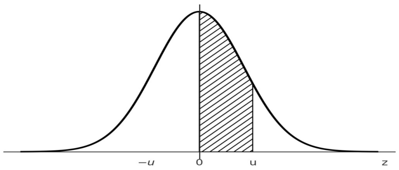


☆そしてこの面積を計算してくれた表を「標準正規分布表」といいます。これを使うと…

問題Ⅰ． 確率変数 Z が正規分布 $N(0,\ 1)$ に従うとき，次の確率を求めよ。

① $P(0\leqq Z\leqq 1.23)$ ＝ 0.3907

標準正規分布に従う確率変数Zの分布曲線

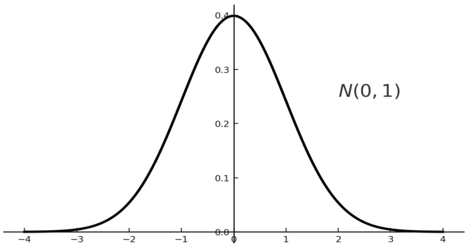


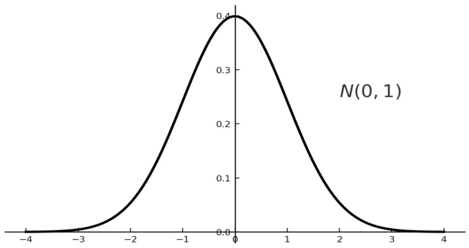
なんと一目で面積すなわち確率が分かってしまう！

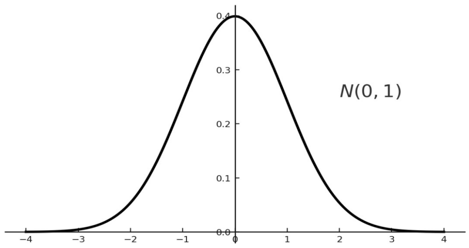
u	……………	.03	
…			
…			
1.2		0.3907	
…			
…			

問題Ⅰ． 確率変数 Z が正規分布 $N(0,\ 1)$ に従うとき，次の確率を求めよ。

② $P(-2\leqq Z\leqq 1.23)$　③ $P(1\leqq Z\leqq 2)$　④ $P(2\leqq Z)$







解いてみた感想と問題２への展望　◎評価ポイント

第６章　「正規分布」⇒「標準正規分布」が出来たら最高！

第４章の問題２は「標準正規分布」ではないからかなり苦しい…。

しかし、実はどんな正規分布でも標準正規分布に変換することができるのです。そんな魔法の式が…

$Point.$ **（正規分布⇒標準正規分布）**

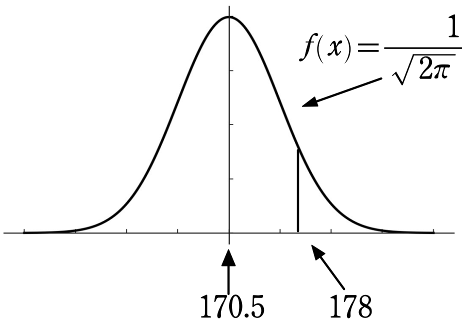
確率変数 X が正規分布 $N(m\,,\ \sigma^2)$ に従う ⟶ 確率変数 Z は標準正規分布 $N(0\,,\ 1)$ に従う

$$\boxed{Z=\frac{X-m}{\sigma}}$$

☆このアイテムがあれば問題２のようなものでも解けるはず！ということで実際に解いてみましょう。

問題２． 確率変数 X が正規分布 $N(170.5\,,\ 5.4^2)$ に従うとき，確率 $P(X \geqq 178)$ を求めよ。

確率変数 X が正規分布 $N(170.5\,,\ 5.4^2)$ に従う



$$f(x)=\frac{1}{\sqrt{2\pi}\cdot 5.4}e^{-\frac{(x-170.5)^2}{2\cdot 5.4^2}}$$

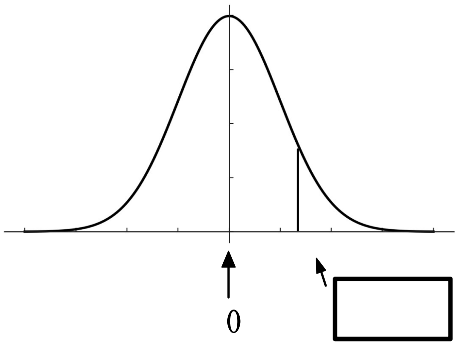
$$P(X \geqq 178)=0.5-P(0 \leqq X < 178)$$

$$=0.5-\int_0^{178}\frac{1}{\sqrt{2\pi}\cdot 5.4}e^{-\frac{(x-170.5)^2}{2\cdot 5.4^2}}\,dx$$

※この計算が苦しくて断念しましたよね。

$$\boxed{Z=\frac{X-170.5}{5.4}}$$

確率変数 Z は標準正規分布 $N(0\,,\ 1)$ に従う



$$P(X \geqq 178)=P(Z \geqq \boxed{}\,)$$

$$=$$

この問題を文章題に変換してみよう　◎評価ポイント

問題３． ある県における高校１年生の男子の身長の平均は 170.5cm，標準偏差は 5.4 cmである。身長の分布を正規分布とみなすとき，次の問いに答えよ。ただし，小数第２位を四捨五入して小数第１位まで求めよ。

(1) 身長 175 cm以上の人は約何%いるか。

(2) 高い方から４%以内の位置にいる人の身長は何cm以上か。

(3) 身長が 160 cm以上 165 cm以下の人は，約何%いるか。

☆標準正規分布の威力は実感できましたか？

ではここまで学んだことを自分の言葉でまとめた上で、最終問題に取り組んでいきましょう。

第１章～第６章までのまとめ　◎評価ポイント

$Point.$ **（二項分布）** これは学習済です

二項分布 $B(n\ ,\ p)$ ⟶ 期 待 値：$E(X)=np$

分　　散：$V(X)=npq$　　※ $q=1-p$

標準偏差：$\sigma(X)=\sqrt{npq}$

※ n：試行の回数，　p：ある事象が起こる確率

☆さてここで問題が起きてきます。今まで正規分布⇒標準正規分布はやってきたのですが、二項分布はやっていません。しかし第１章を思い出してみてください。コインは確かに二項分布に従いますが、結局形は正規分布に近い形になりました。つまり、近似的に正規分布としてみなしてよいということです。すなわち…

$Point.$ **（二項分布⇒正規分布）**

確率変数 X が二項分布 $B(n\ ,\ p)$ に従う ⟶ 確率変数 X は正規分布 $N(\boxed{},\boxed{})$ に従う

☆では１億円をかけた大勝負にいきましょう。

$\boxed{\text{最終問題}}$ 君たちは今１億円を賭けた大勝負のまっただ中にいる。勝負の方法はいたって簡単。

【ルール】１枚のコインを２０回投げて、表が多く出た方の勝ち。イカサマが発覚した時点で負け。

【状　況】相手が先攻で、１５回表が出てしまった。なかなか難しい状況になった…。

君たちの選択肢は以下の２つ。　※確率５％未満は非現実的であると考えてよい。

選択肢①：このまま１６回以上表が出ることを信じて戦う　⇒　勝てる確率はいかほどか

選択肢②：相手のコインの偏りに言及してイカサマを暴きにかかる　⇒　表裏に偏りアリと言えるか

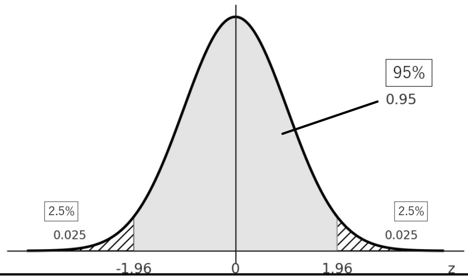
解答

表が出る回数 X は二項分布 $B\left(20\ ,\ \dfrac{1}{2}\right)$ に従っていると仮定する。（コインの裏表に偏りが無いと仮定する）

最終問題を終えて（なんでもＯＫ）　◎評価ポイント

上手く結論までたどり着けたでしょうか？実は確率まで出さなくても、次のことを知っていれば結論まで至れます。

この図の意味するものは？そして何ができるでしょう？　◎評価ポイント



第７章　さあ最後の内容「仮説検定」！

※といってもほぼ終わっています。

Point. **（仮説検定）**

①仮説を立てる

②有意水準 α を定め，棄却域を求める

③確率変数の値が棄却域に入れば仮説を棄却、棄却域に入らなければ仮説は棄却しない

☆難しそうに見えますが、非常に簡単です。最終問題を例にやってみましょう。

最終問題　コイン２０回投げて１５回表が出たけど、このコインって裏表に偏りがあるの？　（簡単に言い換えました）

①仮説を立てよう

②優位水準 α を定め，棄却域を求めよう

$\alpha=$　　　　　，棄却域：

③確率変数を求めて、棄却域に入っているか確かめよう

☆こんな方法でコインの偏りなどを判断することが可能です。

ただ今回は棄却域に入りましたが、棄却域に入らなかったら何が言えるのでしょうか？

もし棄却域に入らなかったら…　◎評価ポイント

☆今回のような検定のやり方を**両側検定**と呼びます。

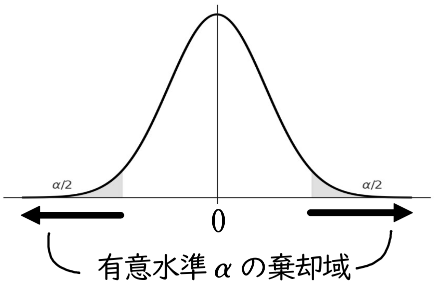
コインのように以上に小さい値や大きい値が出ても「偏りがない」という仮説が棄却されるので両側に棄却域が出ます。

しかし実は**片側検定**というものもあり、次のような問題は片側のみで検定します。

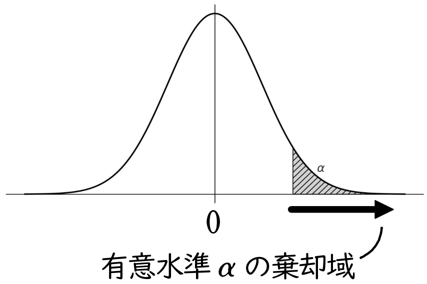
片側検定問題　ある種子の発芽率は従来60％であったが，それを発芽しやすいように品種改良した新しい種子から無作為に150個を抽出して種をまいたところ，102個が発芽した。品種改良によって発芽率が上がったと判断してよいかを，有意水準5％で検定してみよう。

片側のみで良い理由

両側検定



片側検定



☆お疲れさまでした。主に正規分布について話をしてきました。いかがだったでしょうか。

正規分布に対する率直な感想　◎評価ポイント

正規分布になりそうな身近な例　◎評価ポイント

正規分布を利用して他にできそうなこと　◎評価ポイント