

統計的な推測

正規分布⇒仮説検定

教材の方針・使用方法 3～4時間想定，冊子にして配付

【主体的に取り組む態度の育成】

- ①何のために正規分布を学ぶのかを明確にするため、最初に問題を提示する。
- ②そこで起こりうる課題を生徒が発見するよう工夫する。
- ③さまざまなデータを予想し、正規分布を知る。コイン投げのサイトで実際にやるのもよい。**第1～2章**
- ④確率は面積であることの基礎知識を与える。**第3章**
- ⑤正規分布を活用しようとしたときに、計算が困難なことに気付くようにする。**第4章**
- ⑥標準正規分布の有用性を知り、正規分布→標準正規分布にしたいと思うようにする。**第5章**
- ⑦正規分布→標準正規分布の変換方法を知る。**第6章**
- ⑧最初に提示した問題を正規分布を利用して解き、正規分布の有用性を実感する。
- ⑨実はほぼ仮説検定をやっていたことに気付くようにする。**第7章**
- ⑩正規分布の他の利用可能性を考える機会を設ける。→ここで「推定」のことが出てきてくれると非常に有り難い。

【主体的に学習に取り組む態度の評価】

「評価ポイント」の欄を細かく設定することで、「粘り強く取り組もうとしているか」「自らの学びを見直して次につなげようとしているか」を確認し、評価することができる。

- ①問題点・課題点・展望など⇒粘り強く問題解決に意欲的に取り組もうとしているか。
 - ②まとめ・解いてみた感想など⇒他者との意見交換をしながら、学びを見直して次につなげようとしているか。
- ※各評価ポイントに目安が書いてあるが、各学校に合わせてルーブリックを作成してもよい。

【その後の授業展開】：「標本」や「推定」について（習熟度に応じて）

- ①各々が教科書を読み、理解する。
- ②グループ内で理解したことを共有する。
- ③可能であれば、一つのグループが他のグループに解説する時間をとる。

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ある。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 ② を選択

思考

$${}_{20}C_{15}\left(\frac{1}{2}\right)^{15} \cdot \left(\frac{1}{2}\right)^5 + {}_{20}C_{16}\left(\frac{1}{2}\right)^{16} \cdot \left(\frac{1}{2}\right)^4 + \dots + {}_{20}C_{20}\left(\frac{1}{2}\right)^{20}$$

※この式が書けているかをチェックする。

${}_{20}C_{15}\left(\frac{1}{2}\right)^{15} \cdot \left(\frac{1}{2}\right)^5$ しか書けていない生徒に注意する。

結論

※ここは書けていなくてもよい。

問題点・課題点 ◎評価ポイント

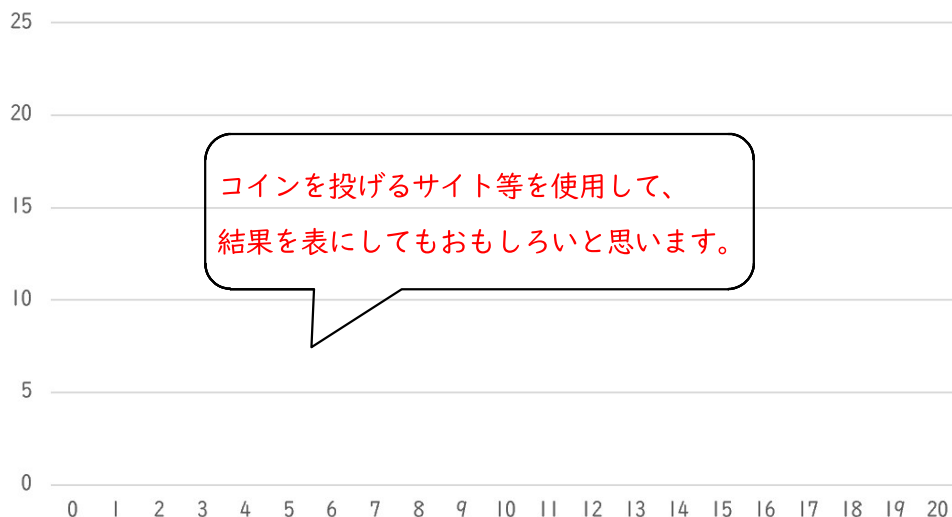
※上記の式を計算することが困難であることに気付く。

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

コインを20回投げて表が出る回数分布(100回の試行)

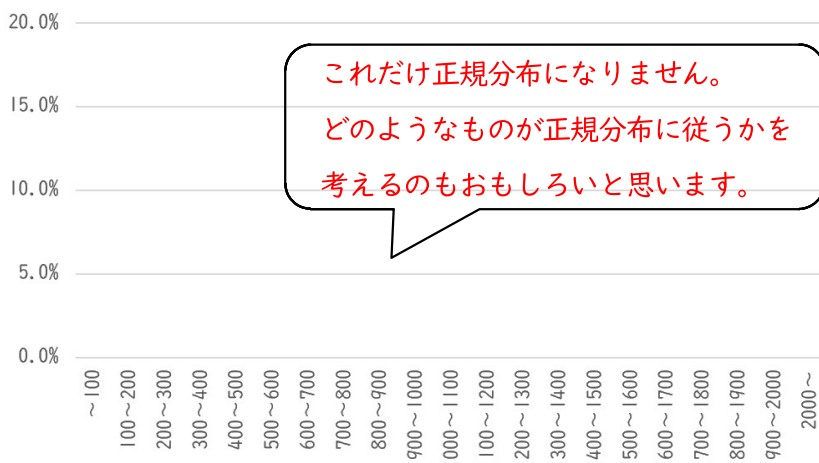


予想②

日本の所得状況の分布

平均値：536万円

中央値：410万円



予想③

17歳日本人男性の身長分布

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。

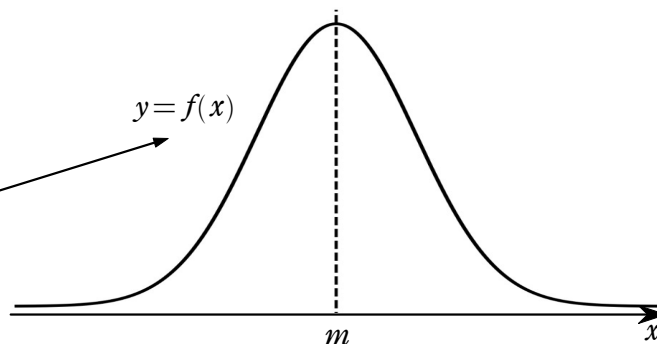
Point. (正規分布)

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！

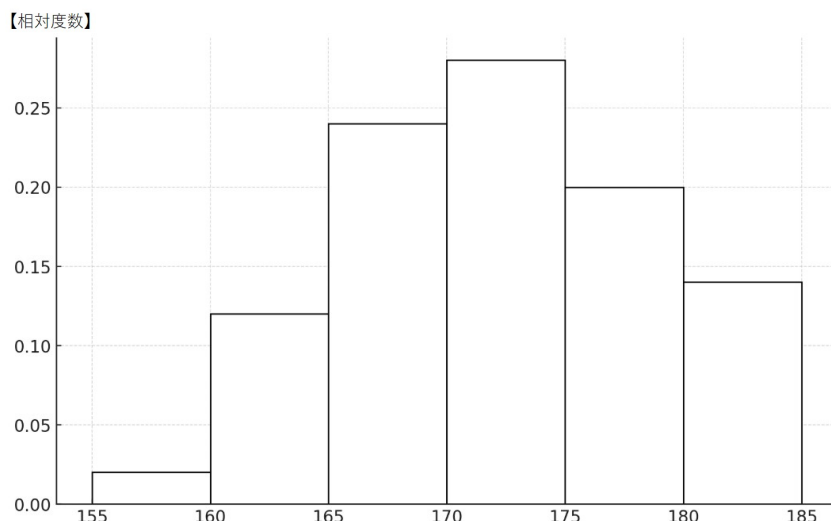


ではこの正規分布の何がすごいのかを理解するために、まずは基礎知識を学んでおきましょう。

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 <small>の階級</small> (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 **確率** 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。

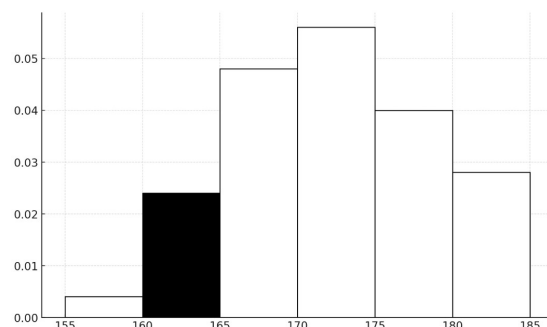
そこでヒストグラム全体の面積を「1」とすると

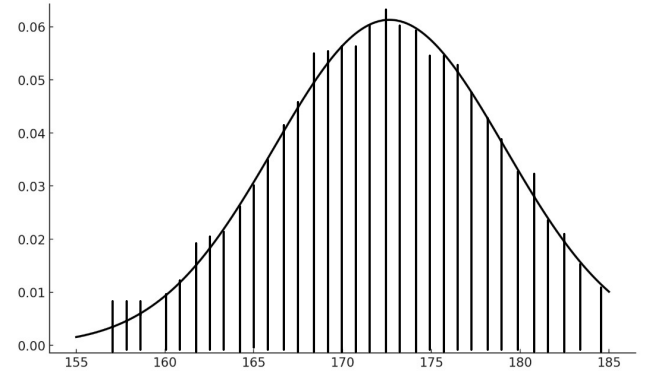
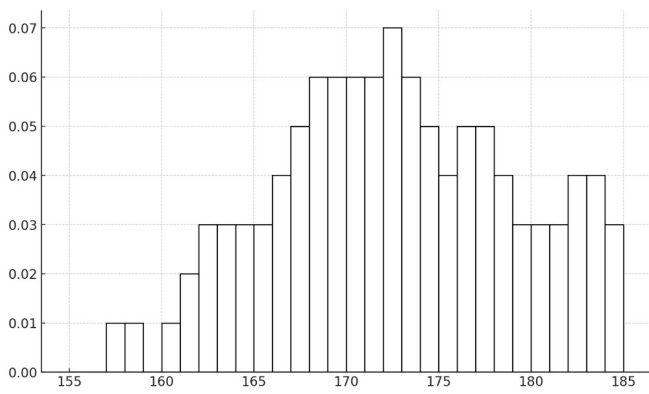
↓

「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)

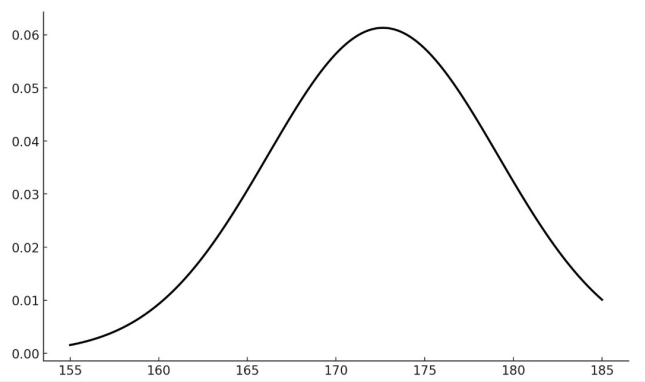
これを確率変数 X を用いて、 $P(160 \leq X \leq 165) = 0.12$ と表現します。

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

$$\int_{170}^{180} f(x) dx$$

ここは立式できればOK。

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

!

「確率の全体=1」を意識させる。

この章をまとめると...

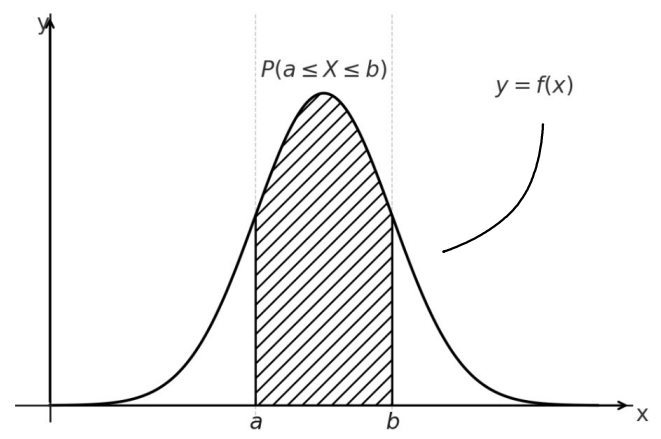
Point. (確率は面積だ!)

① $P(a \leq X \leq b) = \int_a^b f(x) dx$

※確率は面積なので、もちろん積分利用!

② $\alpha \leq X \leq \beta$ のとき $\int_{\alpha}^{\beta} f(x) dx = 1$

※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!

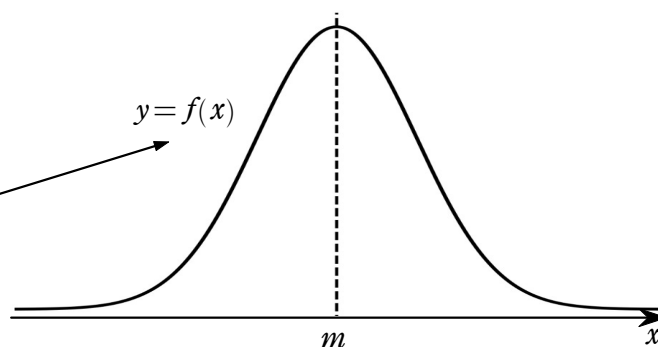
Point. (正規分布)

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 $P(0 \leq Z \leq 1.23)$

$$\int_0^{1.23} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

② $P(-2 \leq Z \leq 1.23)$

$$\int_{-2}^{1.23} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

③ $P(1 \leq Z \leq 2)$

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

④ $P(2 \leq Z)$

$$\int_2^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

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 $P(X \geq 178)$ を求めよ。

$$\int_{178}^{\infty} \frac{1}{\sqrt{2\pi} \cdot 5.4} e^{-\frac{(x-170.5)^2}{2 \cdot 5.4^2}} dx$$

解いてみた感想 ◎評価ポイント

※上記の積分計算が困難であることに気付く。

第5章 「標準正規分布」って最高!

※第3章の考えを忘れずに!

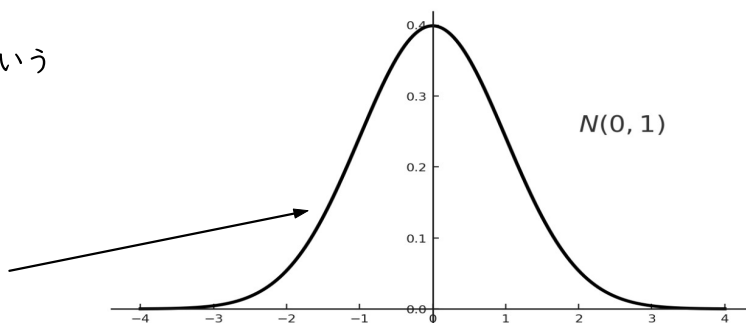
第4章は苦しかったですね。恐ろしく複雑な計算になりました。でも実は…正規分布 $N(0, 1)$ に関してのみ面積を計算してくれた偉大な先人がいるのです。そのような特別な分布のことを次のように呼んでいます。

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…

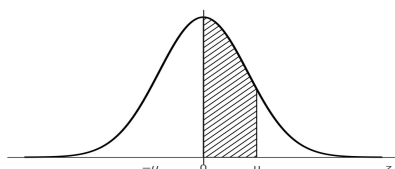
問題1. 確率変数 Z が正規分布 $N(0, 1)$ に従うとき、次の確率を求めよ。

① $P(0 \leq Z \leq 1.23) = \underline{0.3907}$

標準正規分布に従う
確率変数 Z の分布曲線

u	<u>.03</u>
...		↓
...		
<u>.12</u>	→	0.3907
...		
...		

なんと一目で面積すなわち
確率が分かってしまう!

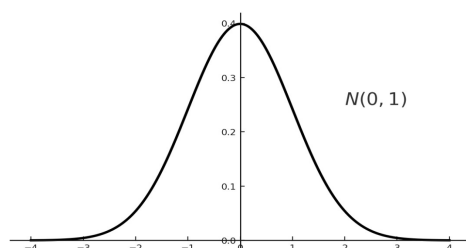
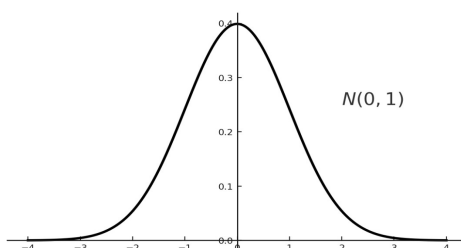
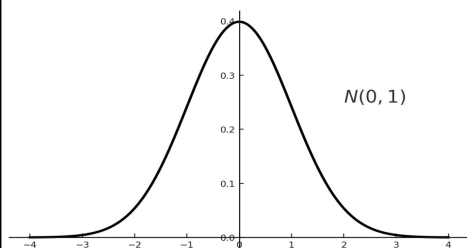


問題1. 確率変数 Z が正規分布 $N(0, 1)$ に従うとき、次の確率を求めよ。

② $P(-2 \leq Z \leq 1.23)$

③ $P(1 \leq Z \leq 2)$

④ $P(2 \leq Z)$



$$\begin{aligned} &P(-2 \leq Z \leq 1.23) \\ &= P(0 \leq Z \leq 2) + P(0 \leq Z \leq 1.23) \\ &= 0.4772 + 0.3907 \\ &= 0.8679 \end{aligned}$$

$$\begin{aligned} &P(1 \leq Z \leq 2) \\ &= P(0 \leq Z \leq 2) - P(0 \leq Z \leq 1) \\ &= 0.4772 - 0.3413 \\ &= 0.1359 \end{aligned}$$

$$\begin{aligned} &P(2 \leq Z) \\ &= 0.5 - P(0 \leq Z \leq 2) \\ &= 0.5 - 0.4772 \\ &= 0.0228 \end{aligned}$$

解いてみた感想と問題2への展望 ◎評価ポイント

※標準正規分布表があって、 $N(0, 1)$ であれば解けることに気付く。

※問題2は $N(0, 1)$ ではないことから、苦しいことに気付く。

※ $N(170.5, 5.4^2) \rightarrow N(0, 1)$ に変えることができないかを考える。

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)

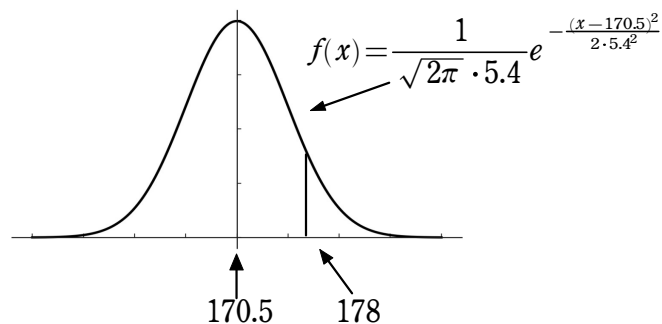
確率変数 X が正規分布 $N(m, \sigma^2)$ に従う \longrightarrow 確率変数 Z は標準正規分布 $N(0, 1)$ に従う

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

☆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 $P(X \geq 178)$ を求めよ。

確率変数 X が正規分布 $N(170.5, 5.4^2)$ に従う



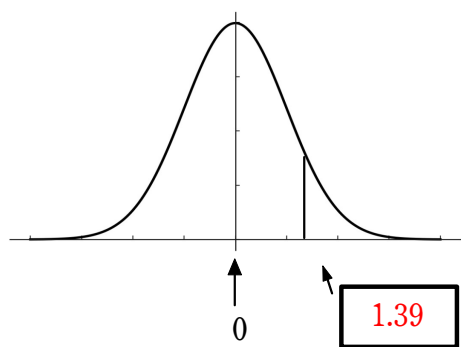
$$P(X \geq 178) = 0.5 - P(0 \leq X < 178)$$

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

※この計算が苦しくて断念しましたよね。

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

確率変数 Z は標準正規分布 $N(0, 1)$ に従う



$$P(X \geq 178) = P(Z \geq \boxed{1.39})$$

=

この問題を文章題に変換してみよう ◎評価ポイント

※身長が平均が170.5cm、標準偏差が5.4の集団において、178cm以上の人が何%いるか?等

※数式⇒問題文への変換を考えることで、自分が今何をやっているのかより理解することができる。

問題3. ある県における高校1年生の男子の身長は平均が170.5cm、標準偏差は5.4 cmである。身長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。

$(1) \frac{175-170.5}{5.4} \doteq 0.65$ <p>よって</p> $P(X \geq 175)$ $= P(Z \geq 0.65)$ $= 0.5 - P(0 \leq Z \leq 0.65)$ $= 0.5 - 0.2422$ $= 0.2578$ <p>以上のことから</p> <p style="border: 1px solid red; padding: 2px; display: inline-block;">約 25.8%</p>	$(2) P(Z \geq u) = 0.04$ $\Leftrightarrow 0.5 - P(0 \leq Z \leq u) = 0.04$ $\Leftrightarrow P(0 \leq Z \leq u) = 0.46$ <p>標準正規分布表より</p> $P(0 \leq Z \leq u) = 1.75$ $1.75 = \frac{X - 170.5}{5.4}$ $\Leftrightarrow X = 179.95$ <p>以上のことから</p> <p style="border: 1px solid red; padding: 2px; display: inline-block;">180.0cm 以上</p>	$(3) \frac{160-170.5}{5.4} \doteq -1.94 \quad \frac{165-170.5}{5.4} \doteq -1.02$ <p>よって</p> $P(160 \leq X \leq 165)$ $= P(-1.94 \leq Z \leq -1.02)$ $= P(1.02 \leq Z \leq 1.94)$ $= P(0 \leq Z \leq 1.94) - P(0 \leq Z \leq 1.02)$ $= 0.4738 - 0.3461$ $= 0.1277$ <p>以上のことから</p> <p style="border: 1px solid red; padding: 2px; display: inline-block;">約 12.8%</p>
---	--	--

☆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

- ※多くのデータが正規分布に従っていることを理解する。
- ※確率を面積で捉えることを理解する。
- ※正規分布に従っているデータを標準正規分布に従うデータに変換できることを理解する。
- ※標準正規分布に持ち込むことで、容易に確率を出すことが可能であることを理解する。
- ※細かいデータがなくとも、平均や標準偏差があれば大体の確率を出せることを理解する。
- ※大きなデータから部分的な割合を出すことが可能であることに気付く。

Point. (二項分布) これは学習済です

二項分布 $B(n, p)$

※ n : 試行の回数, p : ある事象が起こる確率

期待値: $E(X) = np$

分散: $V(X) = npq$ ※ $q = 1 - p$

標準偏差: $\sigma(X) = \sqrt{npq}$

☆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 $N(np, npq)$ に従う

☆では1億円をかけた大勝負にいきましょう。

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ある。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 ⇒ 表裏に偏りありと言えるか

解答

表が出る回数 X は二項分布 $B(20, \frac{1}{2})$ 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)

$$m = 20 \cdot \frac{1}{2} = 10, \quad \sigma^2 = 20 \cdot \frac{1}{2} \cdot \frac{1}{2} = 5 \quad \text{より}$$

近似的に正規分布 $N(10, 5)$ に従う。

よって、

$$Z = \frac{X - 10}{\sqrt{5}} \text{ は近似的に標準正規分布 } N(0, 1) \text{ に従う。}$$

$P(X \geq 15)$ を求める。

$$\frac{15 - 10}{\sqrt{5}} \doteq 2.24 \quad \text{①}$$

これより

$$\begin{aligned} & P(X \geq 15) \\ &= P(Z \geq 2.24) \\ &= 0.5 - P(0 \leq Z \leq 2.24) \\ &= 0.0125 \end{aligned}$$

以上のことから

15回以上表が出る確率は約1.3%であり、表裏に偏りがあると言える。

最終問題を終えて（なんでもOK） ◎評価ポイント

※正規分布の有用性を感じることができているかどうか。

※何をもちて確率5%未満が非現実的と決めるのか。 等

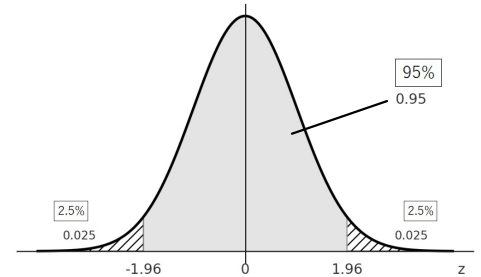
上手く結論までたどり着けたでしょうか？実は確率まで出さなくても、次のことを知っていれば結論まで至れます。

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

※ $Z = \frac{X - m}{\sigma}$ で出てきた値が $Z \leq -1.96, 1.96 \leq Z$ の範囲にあれば

5%未満であることが分かる。

※最終問題波線部①の時点で5%未満が判定できる。等



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

コインの表裏に偏りはない⇒確率はそれぞれ $\frac{1}{2}$

②優位水準 α を定め、棄却域を求めよう

$\alpha = 5\%$, 棄却域: $Z \leq -1.96, 1.96 \leq Z$

③確率変数を求めて、棄却域に入っているか確かめよう

$X = 15$ のとき

$$Z = \frac{15 - 10}{\sqrt{5}} \doteq 2.24$$

これは棄却域である $Z \leq -1.96, 1.96 \leq Z$ に入っているため、仮説は棄却される。

すなわち、コインの裏表に偏りがあると言える。

☆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

※仮説を棄却することはできない。だからといって、仮説が正しいとは限らない。

☆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「偏りが無い」という仮説が棄却されるので両側に棄却域が出ます。

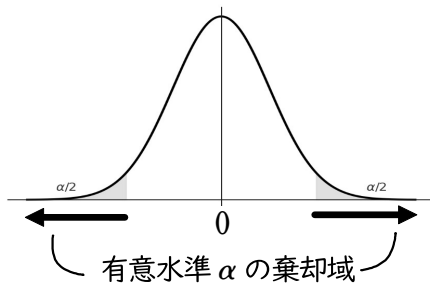
しかし実は**片側検定**というものもあり、次のような問題は片側のみで検定します。

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。

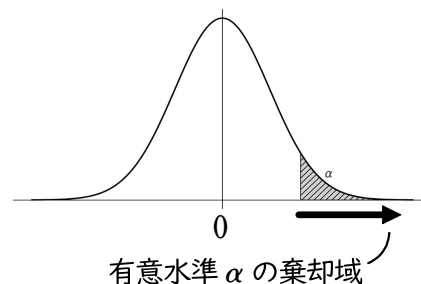
片側のみで良い理由

※発芽率が「上がった」場合のみ考えており、「下がった」場合は考えていないから。

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

※正規分布の有用性を実感している。

正規分布になりそうな身近な例 ◎評価ポイント

※赤ちゃんの出生体重

※テストの得点（偏差値） など

正規分布を利用して他にできそうなこと ◎評価ポイント

※問題3. の逆のイメージで少数のデータから全体のデータを推測できないか。（推定に近いもの）

※偏差値を計算することができる。（データの普通とめずらしいを区別することができる）

※異常なデータを見つけることができる。（工場の製品で平均から大きくずれている⇒不良品の可能性がある） 等