

【資料B】 p 値の出し方を理解しよう

この資料では、 p 値の出し方について学習します。まずは、以下の例題をもとに、 p 値とは何かを説明します。この資料の中で分からない用語があっても、資料Aの担当者が理解しているので安心してください。

例題：コインを20回投げたところ、表が17回、裏が3回出た。このときの p 値を求めなさい。

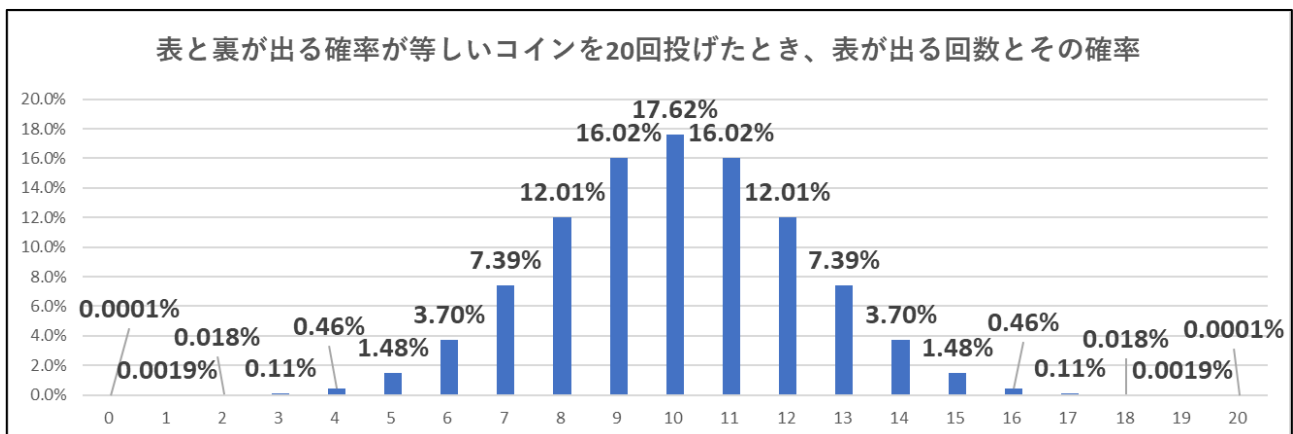
① p 値の定義

p 値は、複数の確率を合計した値です。より具体的には、「帰無仮説が正しいと仮定したときに、統計量の実現値と同じときの確率と、それ以上に極端な値をとるときの確率の合計」と定義されます。「帰無仮説」という言葉は、資料Aの担当者が理解しています。統計量とは、データから求められる色々な数値のことです。そして、実現値とは、統計量の中でも、データの結果に関係する数値のことです。このような説明ではよく分からないと思うので、②で具体的に説明します。

② p 値の計算方法

上記の例題において、「このコインの表と裏が出る確率は等しい」という状況を設定し、「コインを20回投げて表が出る回数」をベースに、 p 値を求めていきます。

さて、コインを20回投げて、表が出る回数とその確率をまとめると、下図のようになります。横軸が「表が出る回数」で、縦軸が確率です。例えば、コインを20回投げたとき表が8回出る確率は、12.01%です。この図をもとに p 値を計算していきます。例題では表が17回出たので、実現値は17です。また今回の場合、表が出る回数の範囲は0～20なので、統計量は、0～20です。



①の p 値の定義に戻ると、 p 値は、「統計量の実現値と同じときの確率+統計量の実現値以上に極端な値をとるときの確率」です。「統計量の実現値と同じとき」、つまり統計量と実現値が 17 のとき、図から確率は 0.11%ということが分かります。「統計量の実現値以上に極端な値をとるとき」、つまり統計量が 18、19、20 のときの確率は、 $0.018+0.0019+0.0001=0.02\%$ です。よって、 p 値は、 $0.11+0.02=0.13\%$ です。かなり低い値が出ました。

③ ツールを使った p 値の出し方

上記のように p 値は計算できます。ですが、実は、js-star というブラウザ上で動く無料ツールを使って、簡単に出すことができます！なので、②の p 値の計算があまり分からなくても、このツールの使い方は習得するようにしましょう。

まずは、「js-star」とネットで検索し、検索結果の最上部の Web サイトを開きましょう。下図の Web サイトが出てくれば OK です。このサイトの左側に、「各種分析ツール」という項目があります。データの種類や統計量によって、 p 値を求めるまでの手順が異なります。そのため、データや統計量によって、この項目から適切なものを選択しなければなりません。今回のように「表/裏」のような 2 択の場合は、「1×2 表(正確二項検定)」と呼ばれる検定方法で p 値を求めます。そのため、左の項目から「1×2 表(正確二項検定)」をクリックしましょう。

★ お知らせ

- TOP
- What's new!
- 動作確認・バグ状況

★ 各種分析ツール

度数の分析

- 1×2 表(正確二項検定)
- 1×2 表：母比率不等
- 1×J 表(カイ二乗検定)
- 2×2 表(Fisher's exact test)

TOP

- すばやいデータ分析を可能にする、ブラウザで動

js-STARは、わかりやすいインターフェースとかんすばやくデータ分析ができる、無償の統計ソフトです
ブラウザ上で動作するため、WindowsでもMacでも作

動作確認は、Windows10 + GoogleChrome で行

- ・第XR版 (js-STAR ver 10) は[こちら](#)です。

すると、右図のようなメニューが表示されます。観測値とは実現値のことです。今回の場合、「表が17回、裏が3回」という実現値が得られているので、観測値は17と3です。これらの値を、「観測値1」と「観測値2」の欄に入力しましょう(順不同です)。どちらも入力すると、観測値の下の「N」の欄に自動で20が表示されます。このNは、観測値の合計です。今回の場合コインを20回投げているのでNは20です。観測値の入力が終わったら、「計算!」をクリックしましょう。これだけで p 値が出力されます。

「計算!」をクリックすると、下図のような結果が表示されます。色々示されていますが、注目するところは、「 $p=0.0013$ (片側確率)」です。この数値は%換算されていないので、%に変換すると0.13%となり、②で計算したものと同じになります。このようにして p 値を出すことができます。

※分析方法によっては、具体的な p 値が出てこないこともあります。以下に出力例を示すので参考にしてください。
「 $p < 0.05$ 」: p 値は1%以上5%未満
「 $p < 0.10$ 」: p 値は5%以上10%未満
「ns」: p 値は10%以上

◆ 【まとめ】問い

「運動部かどうかで一人行動の好き嫌いに差があるか?」を調べるために、100人にアンケートを実施した。その結果が、右図の通りである。このときの p 値は1%より大きいかどうか回答しなさい。

	好き	嫌い
運動部	18	23
運動部以外	44	15

なお、このように2×2表のデータは、正確二項検定ではなくカイ二乗検定という方法で p 値を求める。「各種分析ツール」から「i×j表(カイ二乗検定)」を選択し、表を調整のうえ求めること。

答え: p 値は1%より _____