

そろそろ 統計が面白くなってきた？

いよいよ統計学の本体！

「仮説検定」



統計学は



「できるだけ少ないデータで」

「できるだけ楽をして」

母集団の傾向を知りたい！

という学問

前々回：

「母集団」から「サンプル」を
どれだけ抜き出せば、
母集団の傾向がわかるか調べた！



前回：

「母集団」全データから
母集団の傾向(相関係数)見つけた！

今度は逆から！

今までの2つを使って

「サンプル」は本当に

「母集団の傾向」を表しているのか？

（その仮説が正しいのか）を確認する！



「仮説検定」

「仮説検定」

ある仮説に基づいて行った実験や調査の
今回の結果が「たまたま起こった」だけなのか
について検討する方法



たまたまかどうかの判断には
「確率」を用いる



ただ！　ここで注意！

この仮説検定の中には

「母集団がどんな分布をしているか」
が大事なものがある！

せっ、正規分布？



まずは

「正規分布しているデータ」

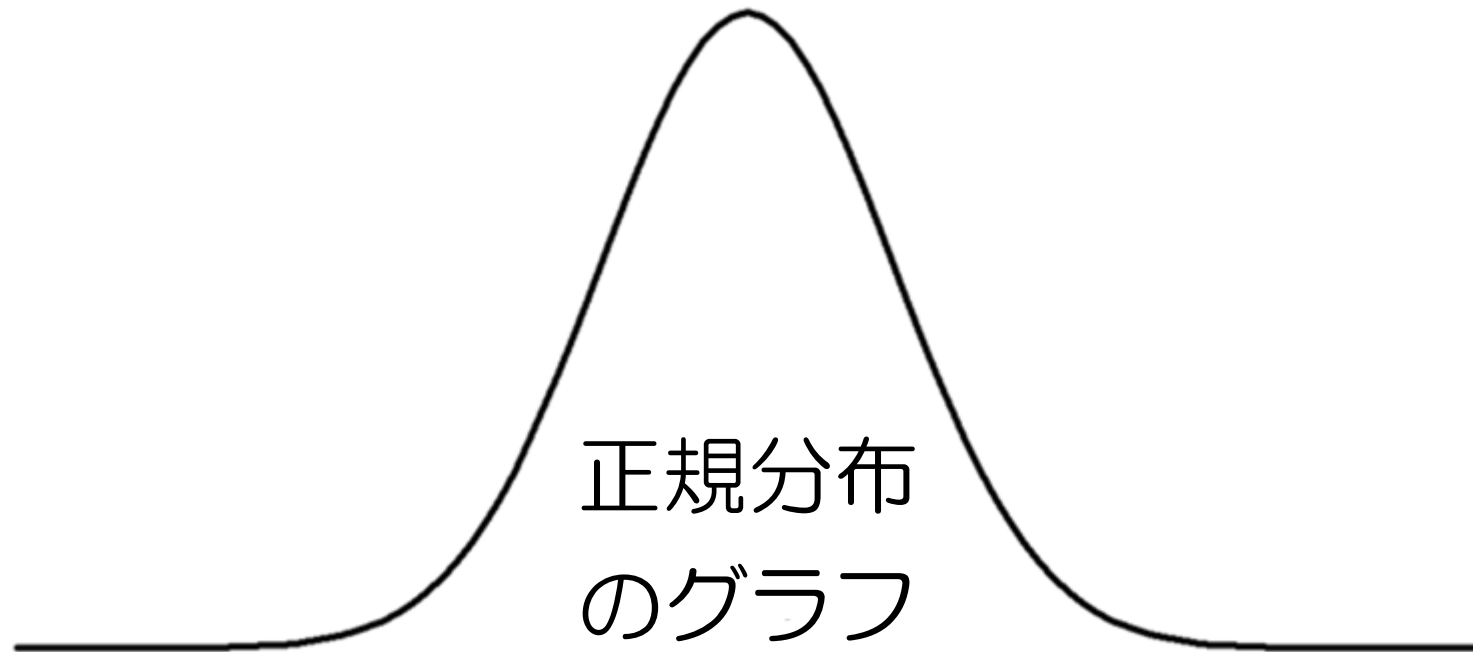
で考えてみよう

「正規分布」 (normal distribution)

「normal」な「ありふれた」分布。

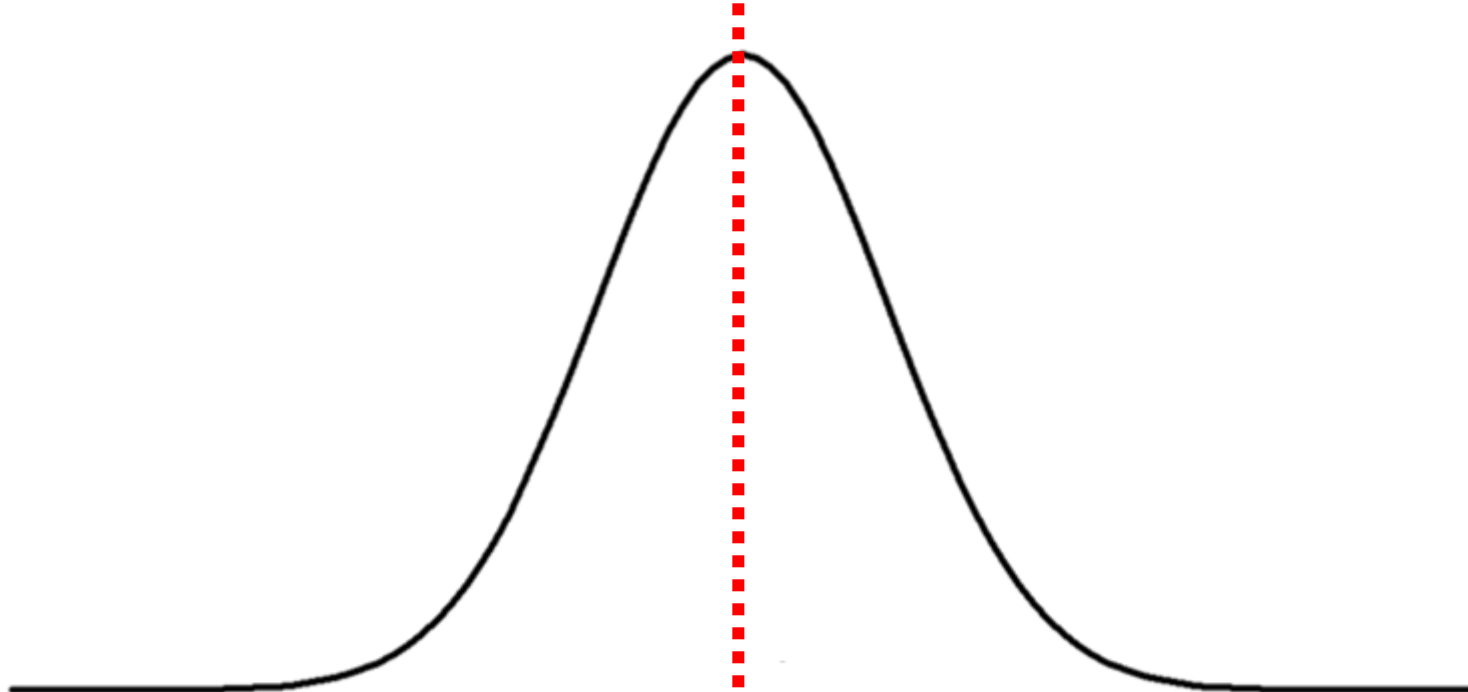
自然界や人間の行動・性質など

いろんな現象に、よく当てはまる



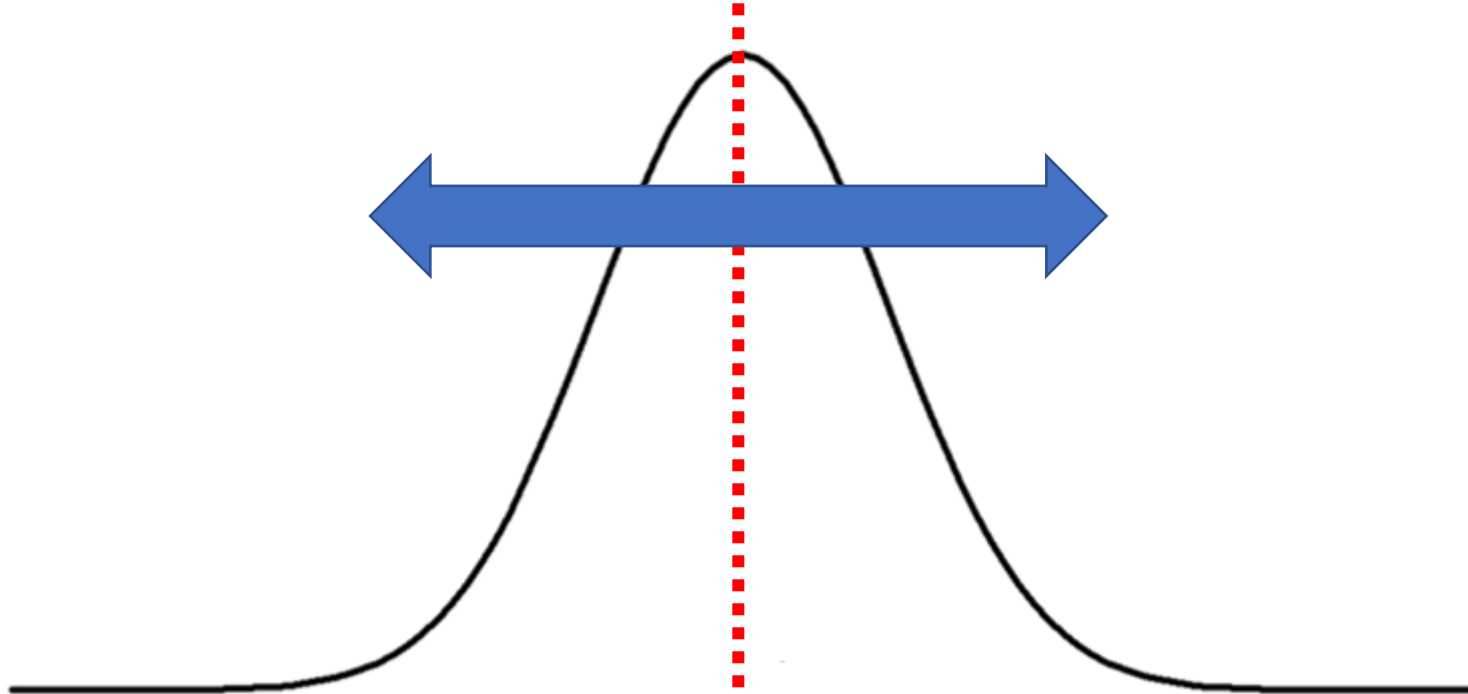
「正規分布」 (normal distribution)

- 平均値と中央値がだいたい一緒
-
-



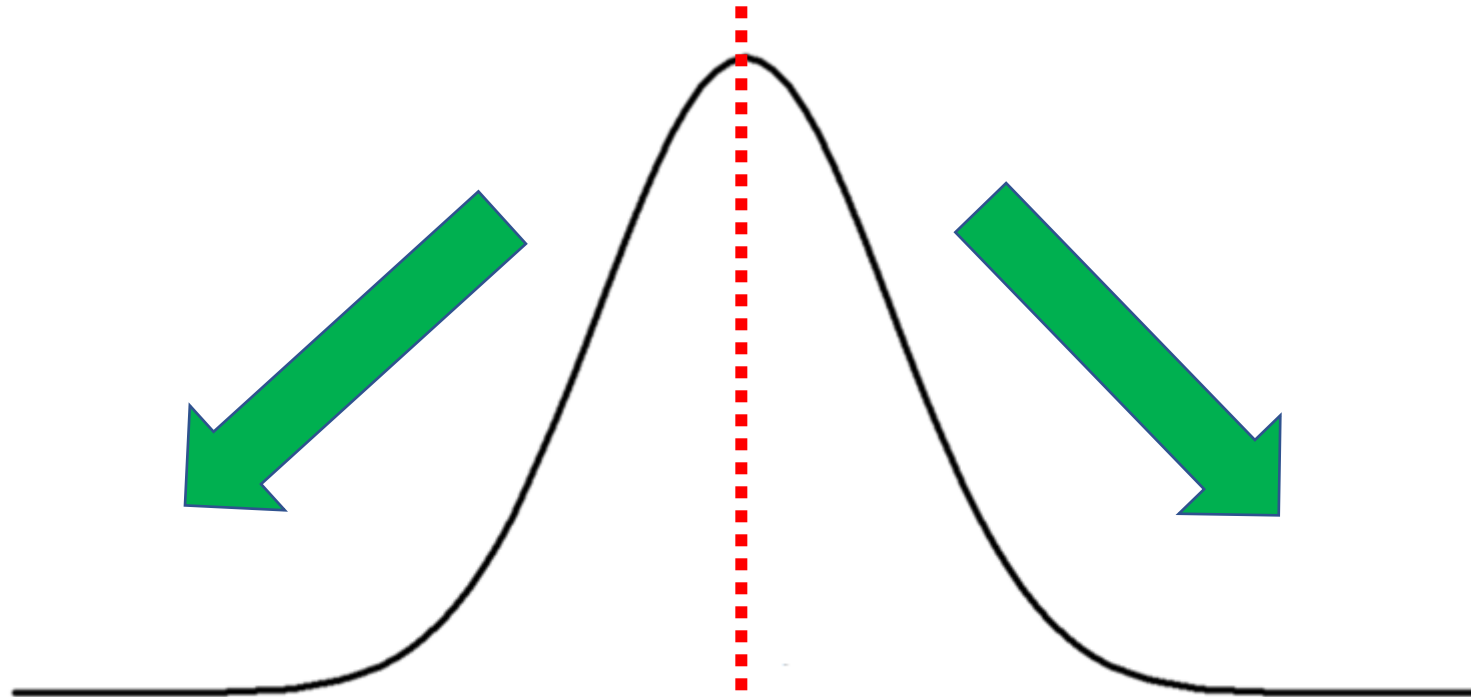
「正規分布」 (normal distribution)

- 平均値と中央値がだいたい一緒
- 平均値を中心にして左右対称



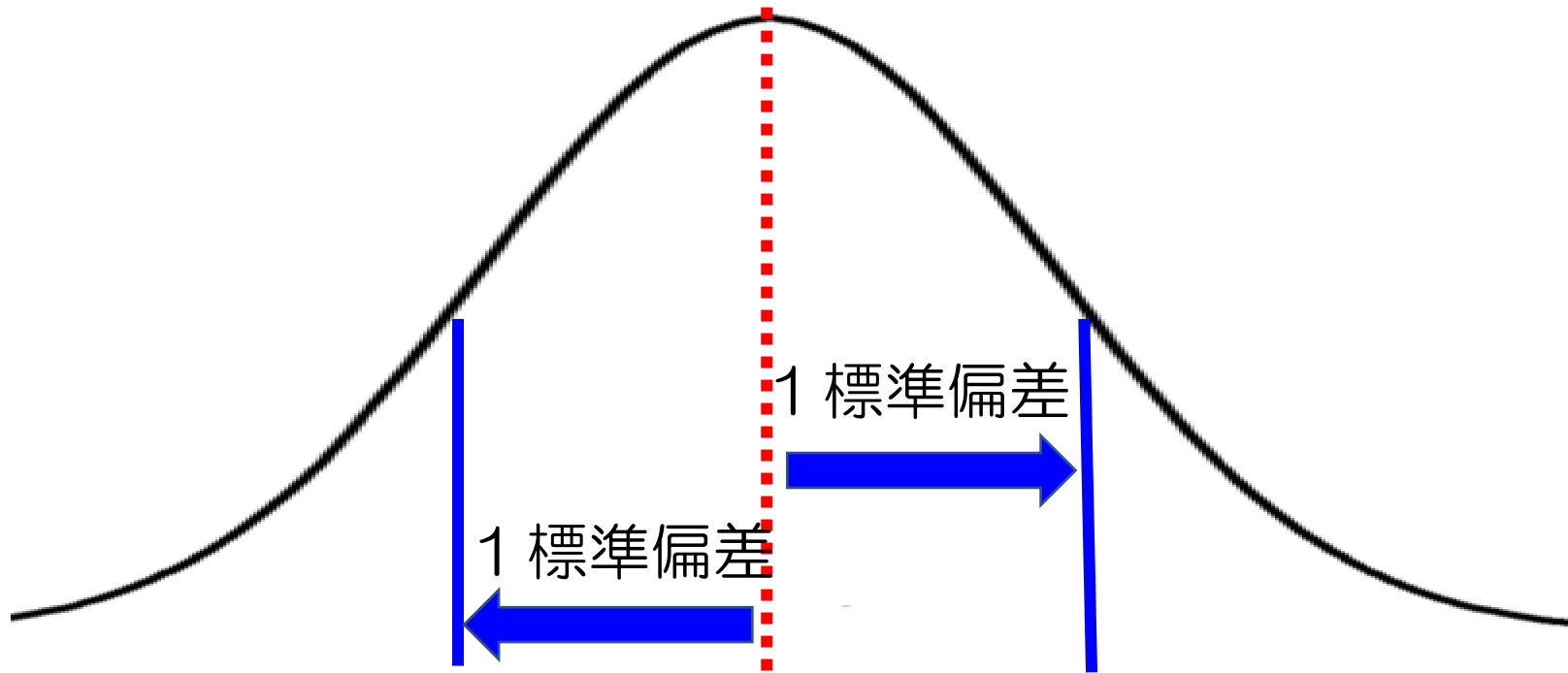
「正規分布」 (normal distribution)

- 平均値と中央値がだいたい一緒
- 平均値を中心にして左右対称
- 平均値から離れるほど少なくなる



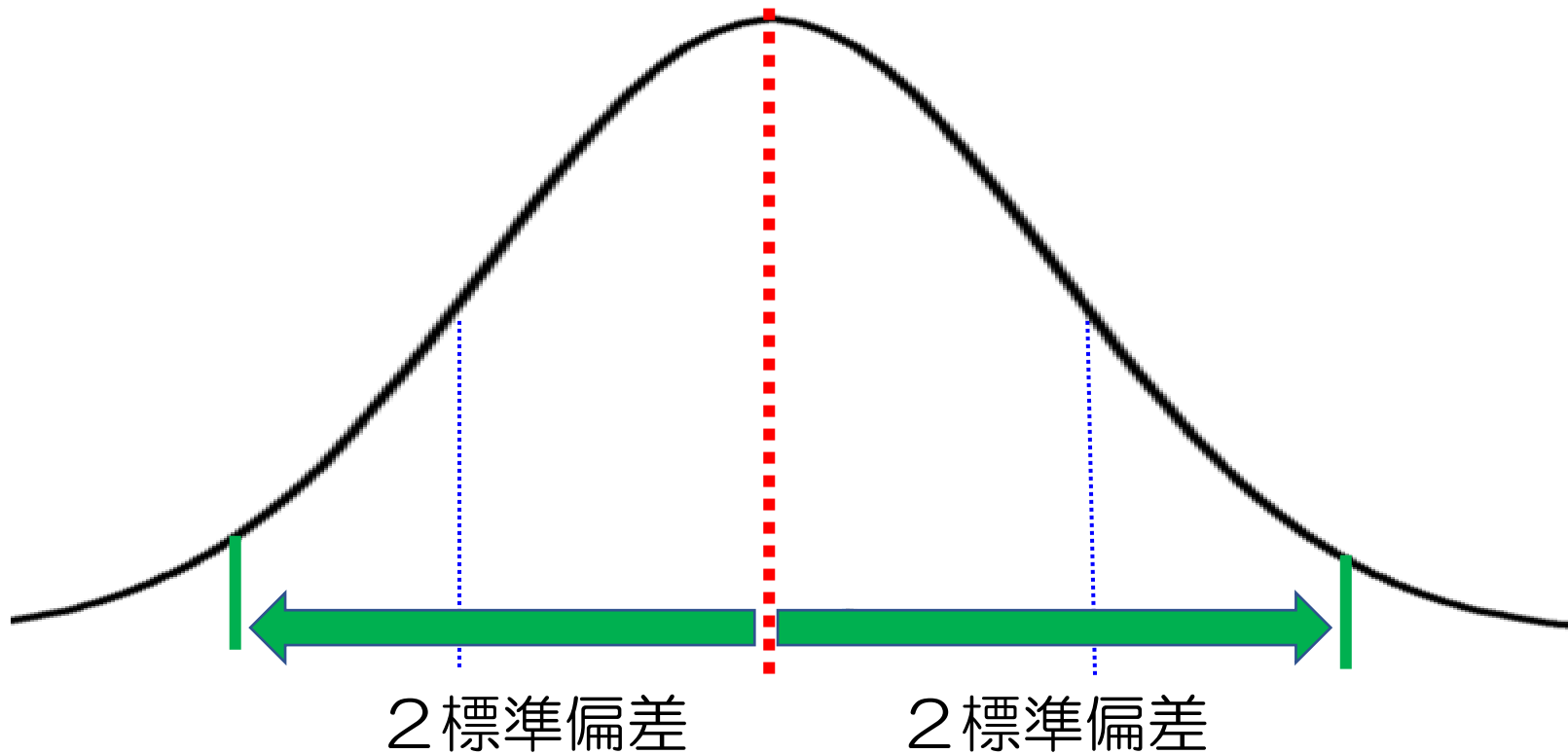
「正規分布 と 標準偏差」

- 平均値 $\pm$ 1 標準偏差の中に
データの「68%」がある



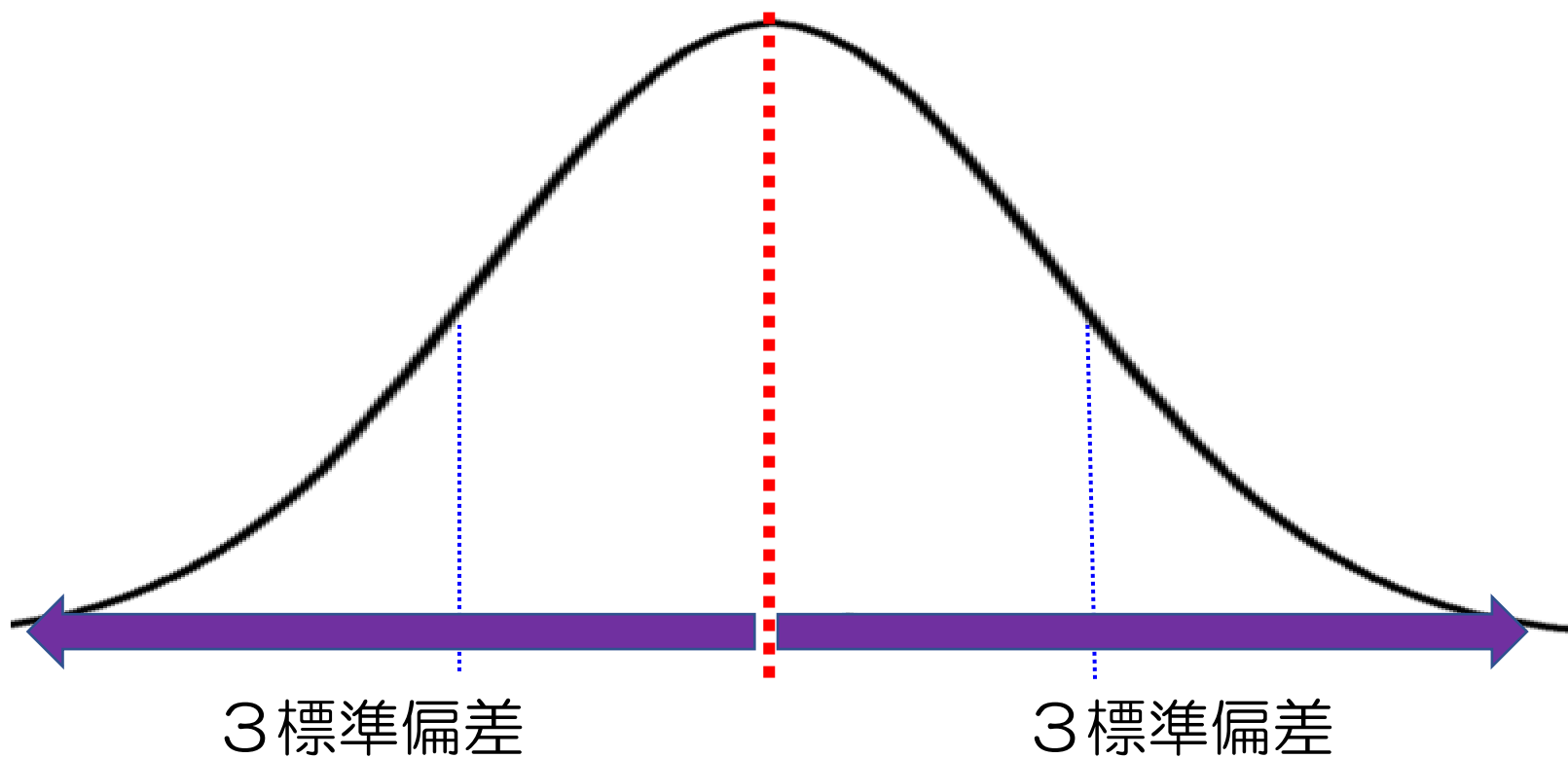
「正規分布 と 標準偏差」

- 平均値 $\pm$ 2標準偏差の中に
データの「95%」がある



「正規分布 と 標準偏差」

- 平均値 $\pm$ 3 標準偏差の中に
データの「99.7%」がある



いろいろな統計的な分析方法

1 カイ二乗検定

2 t 検定

3 回帰分析



天王寺動物園のトラが

「阪神 対 巨人」戦を

予想して当てた。



WIN!



このトラが、
5回連続で
予想を当てたとする



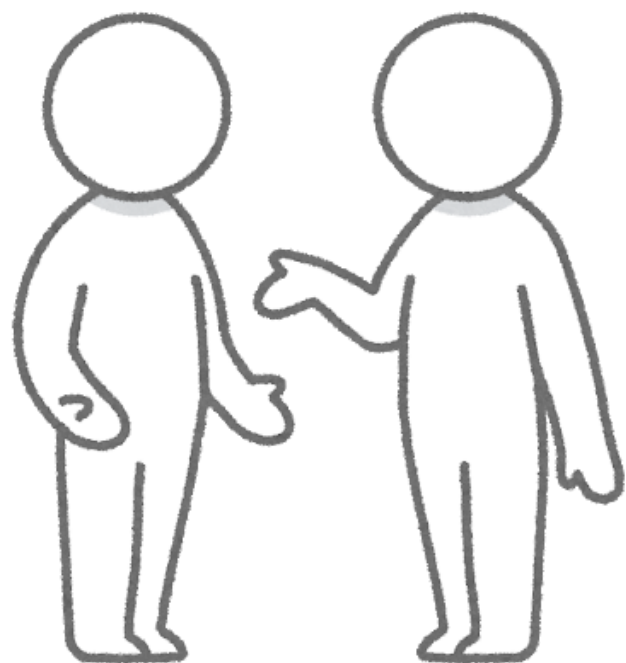
このトラの予想力は本物といえる？

(※試合の勝率は常に $1/2$ と仮定)

Aくん 「5回連続で当ててるから本物！」

Bくん 「5回連続はたまたまやろ！」

Cくん 「次はハズれる」



これを統計学を使って
客観的に判断するのも

「仮説検定」

「仮説検定」

軽く見てみよう



Aくん 「5回連続で当ててるから本物！」

トラの予想力は本物や！

って言いたい！ そのために…



「トラには予想力がある」

と言えるだけの証拠があるか調べたい！

「トラには予想力がある」

と言えるだけの証拠があるか調べたい！



まずは

H_1 : トラには予想力がある

という仮説をたてる



でも…

「トラには予想力がある」
と言えるだけの証拠

← 難しい！！

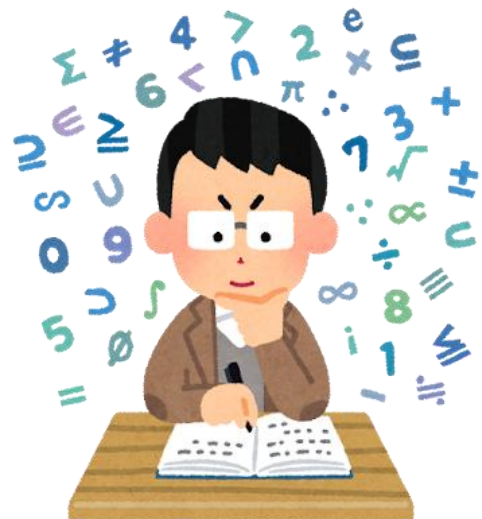


そこで逆に、

「予想力がない」としたら

5連続正解がどれくらい珍しいのか

考える！



「トラには予想力がある」

と言えるだけの証拠があるか調べたい！



なので、次に

H_0 : トラに予想力はない！

という**仮説**をたてる

必ず、自分が主張したいことと
「逆の仮説」を立てる！



今、立てた仮説

対立仮説

H_1 : トラには予想力がある

H_0 : トラに予想力はない！

帰無仮説

対立仮説？

帰無仮説？

よくわからん



例えば

帰ったら、自分の机の上に

「プリン」があった。

(スプーンも)



仮説① H_1 : 誰かが~~買~~ってきてくれた

仮説② H_0 : 風で飛んできた



仮説②は、仮説①を否定する仮説

つまり、仮説①を無に帰する仮説

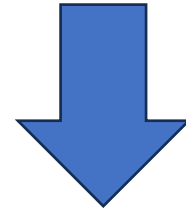
「帰無仮説」



ここで前回までの復習

ほんとは

世界の全員のデータを取りたい！



お金も、時間もかかるし、

実際には無理！

ここで前回までの復習



ランダム抽出

母集団

統計学的仮説検定

標本

実験・調査

実験・調査

母集団の結果

推測

標本からの結果

ランダムでサンプリングするから



母集団からの
結果



標本から
の結果



標本から
の結果



標本から
の結果



標本から
の結果



母集団の結果と
ずれてるんちゃう？





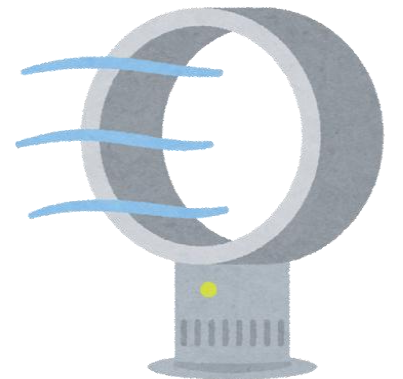
H_1 : 誰かが買ってきてくれた

H_0 : 風で飛んできた

例えば、 H_0 が正しいと仮定したときに

今回の仮説（風で飛んできた）

が起こる確率はどんくらい？





H_1 : 誰かが買ってきてくれた

H_0 : 風で飛んできた

じゃあ、どれくらい珍しかったら
(確率が低かったら)

「 H_0 はあり得ない! (珍しい) こと」って言える?

その判断ラインが 「有意水準 α 」

話は、戻って

H_0 : トラに予想力はない!

帰無仮説



これが間違っていると証明できれば

「トラには予想力がある」

と言える



ってことで、帰無仮説が正しいとして計算してみる

「仮説検定」

H_0 : トラに予想

こいつを「p値」
って呼ぶ！

1/2を5回当てる確率は

$$\left(\frac{1}{2}\right)^5 = 0.0312$$

3.1%

3.1%ってことは、ほぼ有り得ないから、
偶然では説明しにくいんじゃない？

「仮説検定の有意水準」

統計学仮説検定では

有意に（明らかに）間違っていると

判断する基準を確率で表す！

それが「有意水準」(α)

普通は「5%」か「1%」

先に決めておく

「p 値」

H_0 が正しいとしたときに

極端な結果が出る確率

(今回だと3.1%)

H_0 が間違っていると判断するには

「p 値が有意水準 (α) より小さい」

ことが条件！

「有意水準（ α ）と p 値」

有意水準（ α ）	P値（p-value）
偶然の誤差とみなす限界の確率 （合格ライン）	帰無仮説が正しいと仮定した時に、 極端なデータが得られる確率
検定を行う前にあらかじめ決める	データを収集後に計算で決まる
5%（0.05）または 1%（0.01）	データごとに異なる数値

「仮説検定の有意水準」

「有意水準」(α)

普通は「5%」か「1%」

$$\alpha = 0.05$$

$$\alpha = 0.01$$

「5%」：

「 H_0 を棄却する」

$p = 3.1\%$

ということ？

「仮説検定」 最終

$\alpha = 5\%$

$p = 3.1\%$



H_0 : トラに~~予想~~力はない！

H_1 : トラに予想力がある！

p 値が有意水準 α より小さい

時は H_0 を却下できる！

「仮説検定」 最終

$\alpha = 1\%$

H_0 : トラに予想力はない！

H_1 : トラに予想力がある！

$p = 3.1\%$

p 値が有意水準 α より大きい

時は H_0 を却下できない！

どっちが正しいか決められない！



「仮説検定」 最終

H_0 : トラに予想力はない！
 H_1 : トラに予想力がある！

有意水準
5%
(危険率)

H_0 : 「トラに予想力はない！」は棄却できる
つまり、トラには予想力がある！

有意水準
1%

H_0 : 「トラに予想力はない！」は棄却できない
つまり、トラに予想力があるのかわからない



逆に、「5%」のとき

「**棄却**」してよかったのかな？

間違ってたらどうしょ💧

たまには間違ふこともあるよ！

それが「**統計学**」！



「過誤」

H_0 を棄却

判断	H_0 正しい	H_1 正しい
する	第1種の過誤	正しい
しない	正しい	第2種の過誤

第1種の過誤 （ α 過誤：あわてものの誤り）

仮説が正しいのに、棄却してしまった

第2種の過誤 （ β 過誤：ぼんやりものの誤り）

仮説が間違ってるのに、棄却しなかった

「仮説検定」スタート



「仮説検定」の基本的な考え方



「たぶん、こうなるはずや！」

「そんなはずはない！」



これを判断するには、自分の中に**判断基準**が必要
判断基準があるから「ありえない」と感じる！

「仮説検定」の基本的な考え方



「たぶん、こうなるはずや！」

「そんなはずはない！」



その判断基準こそが「有意水準」

「仮説検定の有意水準」

ある仮説が有意に（明らかに）間違っていると
判断する基準を確率で表す！

「有意水準」(α) 普通「5%」か「1%」

なんかピンとこないなあ…



例えば、ある仮説

「今日は傘を持って出かけたほうがいい！」

降水確率が何%なら

傘を持っていく？



降水確率

80%なら

持っていく

50%なら

持っていくかも

20%なら

悩む？

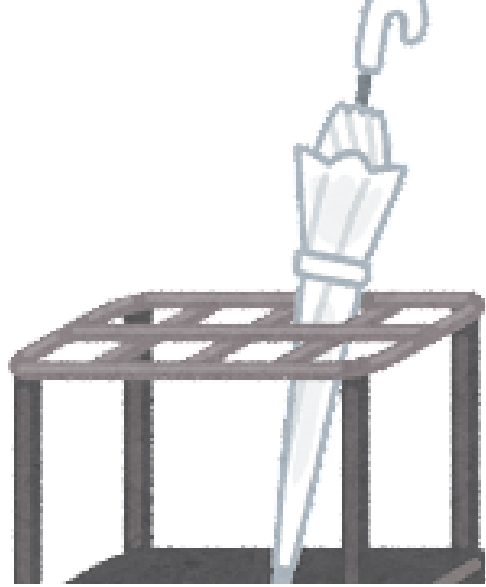
3%なら

？？？

1%なら

？？？





降水確率

80%なら

持っていく

50%なら

持っていくかも

20%なら

悩む？

3%なら

???

1%なら

???



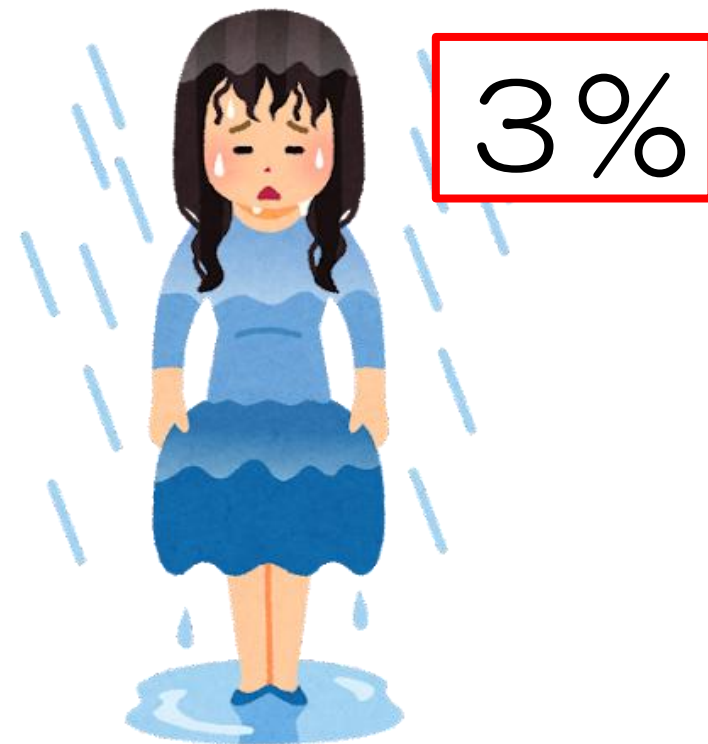
降水確率が3%だから、傘を持ってこなかった
つまり、「雨は降らない」と判断した

降水確率が3%だから、傘を持ってこなかった
つまり、「雨は降らない」と判断した

逆に言うと

3%の可能性で

雨に濡れる危険がある！



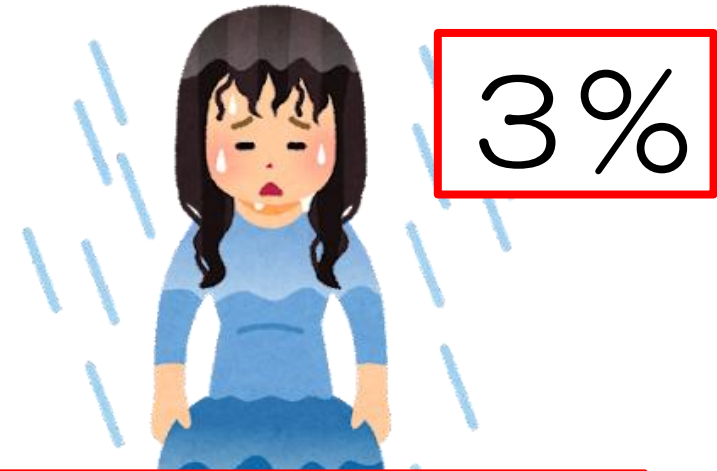
「今日は傘を持って出かけたほうがいい！」

という仮説に対して

「3%」の危険をおかしても、

仮説が間違っていると思うから

傘を持ってこなかった



ある仮説が有意に（明らかに）間違っていると
判断する基準を確率で表す！

天気予報の降水確率が「3%」のとき

自分の判断の基準になる降水確率が

「5%」なら 傘を持っていかない

「1%」なら 傘をもっていく

3%



「有意水準」(α) 普通「5%」か「1%」

「仮説検定」の種類



いろいろな統計的な分析方法

1 カイ二乗検定

2 t 検定

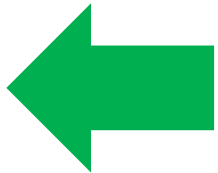
3 回帰分析



	結果が2択 「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを 使う？ 使う？使 わない？ 2択	×二乗検定	t 検定
量はどれ くらい？ どれくら い程度？ 複数	ロジスティック回帰	回帰分析

	結果が2択 「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを使う？ 使う？使わない？ 2択	説明変数 独立変数	↑ 目的変数 従属変数
量はどれくらい？ どれくらい程度？ 複数		

	結果が2択 「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを使う？ 使う？使わない？ 2択	← タバコ吸う？・吸わない？ 放射線治療する？・しない？ 朝食食べてる？・食べない？	
量はどれくらい？ どれくらい程度？ 複数		

	結果が2択 「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを 使う？ 使う？使 わない？ 2択		
量はどれ くらい？ どれくら い程度？ 複数	 タバコ何本吸う？ 放射線どれくらいあてる？ 朝食何カロリーくらい？	

	結果が2択「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを 使う？ 使う？使 わない？ 2択	↑ 血圧上がった？下がった？ がん消えた？消えてない？ テスト合格？不合格？	
量はどれ くらい？ どれくら い程度？ 複数		

	結果が2択 「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを 使う？ 使う？使 わない？ 2択		 血圧はいくつになった？ がんはどれくらい消えた？ テスト何点やった？
量はどれ くらい？ どれくら い程度？ 複数		

	結果が2択「○・×」 「合格不合格」「あり・なし」	結果は程度（複数） 「数値」「点数」「金額」とか
どっちを使う？ 使う？使わない？ 2択	↑結果は変わる？ ← 結果 これによって 原因	
量はどれくらい？ どれくらい程度？ 複数		

結果が2択

×二乗検定

条件
2択

朝食○?×?

合格?不合格?

朝食を食べるか食べないかで
テストの合格不合格に差があるのか?

		結果は複数
条件 2択		t 検定 朝食○?×? テスト点数に差?
	朝食を食べるか食べないかで テストの点数に差があるのか?	

結果が2択

何km走れば
ダイエットは成功・失敗するのか？

条件
複数

何キロ走れば

ダイエットできる？
できない？

ロジスティック回帰

		結果は複数
	何km走れば 体重はどれくらい減るのか？	
条件 複数		何キロ走れば どれくらい減る？ 回帰分析

いろいろな統計的な分析方法

1 カイ二乗検定

2 t 検定

3 回帰分析



結果が2択

×二乗検定

条件
2択

朝食○?×?

合格?不合格?

朝食を食べるか食べないかで
テストの合格不合格に差があるのか?

「×二乗検定」

朝食を 「食べる 食べない」と

テストに「合格 不合格」 に関係がある？

「2択×2択」か…

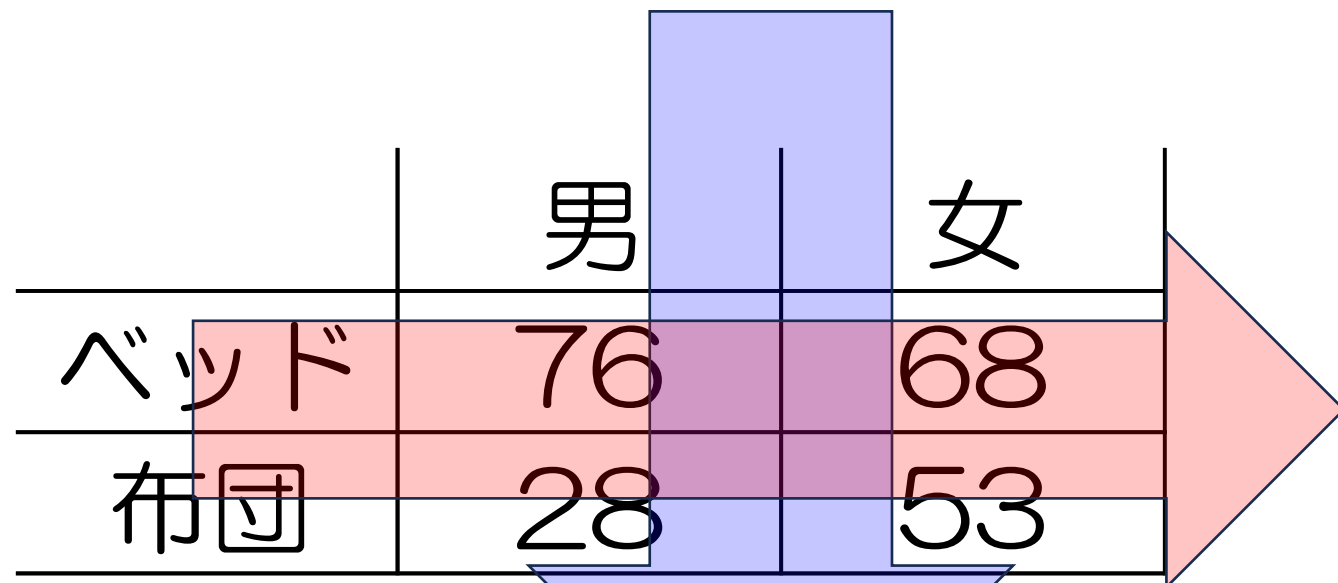
どっかでみた気がする

「クロス集計表」



「クロス集計表」

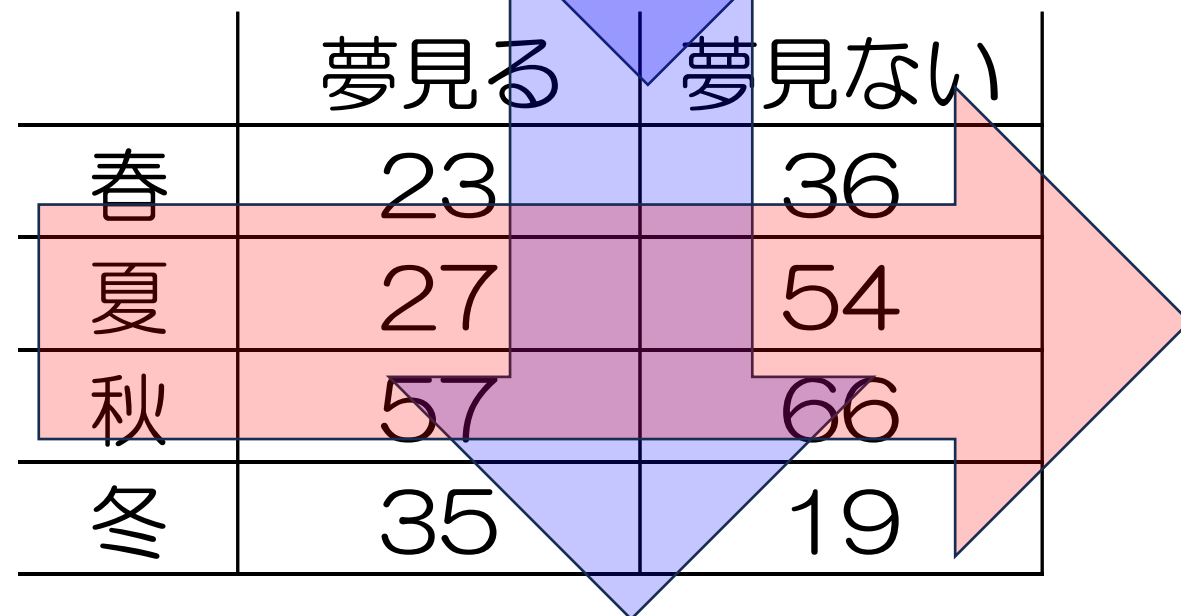
前回やったやつ！



A cross-tabulation table showing the relationship between Bedding (ベッド) and Gender (性別). The table is presented in a butterfly format, with male data on the left and female data on the right. The central column is shaded blue. The rows are shaded pink. A large pink arrow points to the right on the far right side.

	男	女
ベッド	76	68
布団	28	53

複数でもできる！



A cross-tabulation table showing the relationship between Dreaming (夢見る) and Season (季節). The table is presented in a butterfly format, with 'dreaming' data on the left and 'not dreaming' data on the right. The central column is shaded blue. The rows are shaded pink. A large pink arrow points to the right on the far right side.

	夢見る	夢見ない
春	23	36
夏	27	54
秋	57	66
冬	35	19

例えば、「赤いお皿」と「青いお皿」

どっちかを自由にとってもらう



取ったお皿のお菓子を

「食べるか」「食べないか」 観察



「お皿の色とお菓子を食べるか

どうかには関係がある」と仮説を立てて検定

Excel シート

「お皿の色ー食べたかどうか」

「提出 検定結果」例のところ

まずは、こっちを見て！



「 χ^2 乗検定」

H_0 「お皿の色と食べるかは関係がない」

H_1 「お皿の色と食べるかは関係がある」

と仮説を立てて検定

H_1 に「自分が証明したい仮説」をたてる！

Excel シート

「お皿の色ー食べたかどうか」

「提出 検定結果」例のところ



次は、こっちを見ながら！

χ²乗検定の流れ

- 1 データをクロス集計表（2×2）にまとめる
- 2 期待度数を求める
（もし関係が無かったら、こうなるやろという値）
- 3 データと期待度数との差を求める（χ²乗の準備）
（この差が大きければ、関係がありそう）
- 4 χ²乗値を求める
- 5 χ²乗値をp値に変換する
- 6 p値から判断する

χ²乗検定の流れ

1 データをクロス集計表（2×2）にまとめる

2 期待度数を求める

（もし関係が無かったら、こうなるやろという値）

3 データと期待度数との差を求める（χ²乗の準備）

（この差が大きければ、関係がありそう）

4 χ²乗値を求める

5 χ²乗値をp値に変換する

6 p値から判断する

1 データをクロス集計表にまとめる

赤いお皿を取って食べた人 70人
食べなかった人 180人

青いお皿を取って食べた人 30人
食べなかった人 120人

	食べた	食べてない	合計
赤	70	180	250
青	30	120	150
合計	100	300	400

χ²乗検定の流れ

1 データをクロス集計表（2×2）にまとめる

2 期待度数を求める

（もし関係が無かったら、こうなるやろという値）

3 データと期待度数との差を求める（χ²乗の準備）

（この差が大きければ、関係がありそう）

4 χ²乗値を求める

5 χ²乗値をp値に変換する

6 p値から判断する

2 期待度数を求める

(もし関係が無かったら、
こうなるやろという値)

	食べた	食べてない	合計
赤	62.5	187.5	250
青	37.5	112.5	150
合計	100	300	400

2 期待度数を求める

	食べた	食べてない	合計
赤	70	180	250
青	30	120	150
合計	100	300	400

全体の400人のうち食べた人は「100人」

ということは、

何も関係が無ければ「1/4」は食べるはず！

2 期待度数を求める

	食べた	食べてない	合計
赤	70	180	250
青	30	120	150
合計	100	300	400

赤いお皿を取った人は「250人」

2 期待度数を求める

	食べた	食べてない	合計
赤	70	180	250
青	30	120	150
合計	100	300	400

関係が無ければ「1/4」は食べるはず！

$$250 \times 1/4 = 62.5$$

期待値	食べた	食べてない	合計
赤	62.5		250
青			150
合計	100	300	400

2 期待度数を求める

	食べた	食べてない	合計
赤	70	180	250
青	30	120	150
合計	100	300	400

関係が無ければ「3/4」は食べない！

$$250 \times 3/4 = 187.5$$

期待値	食べた	食べてない	合計
赤	62.5	187.5	250
青			150
合計	100	300	400

2 期待度数を求める

期待値	食べた	食べてない	合計
赤	62.5	187.5	250
青	37.5	112.5	150
合計	100	300	400

同じようにもとめる

$$150 \times 1/4 = 37.5$$

$$150 \times 3/4 = 112.5$$

χ²乗検定の流れ

1 データをクロス集計表（2×2）にまとめる

2 期待度数を求める

（もし関係が無かったら、こうなるやろという値）

3 データと期待度数との差を求める（χ²乗の準備）

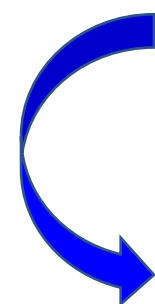
（この差が大きければ、関係がありそう）

4 χ²乗値を求める

5 χ²乗値をp値に変換する

6 p値から判断する

3 データと期待度数との差を求める (χ二乗の準備)



データ値		食べた	食べてない	合計
	赤	70	180	250
	青	30	120	150
	合計	100	300	400

期待度数		食べた	食べてない	合計
	赤	62.5	187.5	250
	青	37.5	112.5	150
	合計	100	300	400

引き算するだけ

χ二乗 準備		食べた	食べてない	合計
	赤	7.5	-7.5	
	青	-7.5	7.5	
	合計			

3 データと期待度数との差を求める（ χ^2 乗の準備）

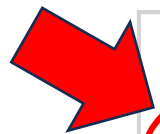
「 χ^2 乗の準備」

別にこういう値があるわけじゃないよ！

この次にやる計算が楽になると思うから

やってもらってるだけ！



 χ^2 乗 準備		食べた	食べてない	合計
	赤	7.5	-7.5	
	青	-7.5	7.5	
	合計			

χ²二乗検定の流れ

- 1 データをクロス集計表（2×2）にまとめる
- 2 期待度数を求める
（もし関係が無かったら、こうなるやろという値）
- 3 データと期待度数との差を求める（χ²二乗の準備）
（この差が大きければ、関係がありそう）
- 4 χ²二乗値を求める
- 5 χ²二乗値をp値に変換する
- 6 p値から判断する

4 X二乗値を求める 「X二乗準備の値」

(データの値ー期待度数)²

期待度数

を計算する

		食べた	食べてない	合計
X二乗	赤	0.9	0.3	
	青	1.5	0.5	
	合計			

この4つを足したものが「X二乗値」
 $0.9 + 0.3 + 1.5 + 0.5 = 3.2$

χ²乗検定の流れ

- 1 データをクロス集計表（2×2）にまとめる
- 2 期待度数を求める
（もし関係が無かったら、こうなるやろという値）
- 3 データと期待度数との差を求める（χ²乗の準備）
（この差が大きければ、関係がありそう）
- 4 χ²乗値を求める
- 5 χ²乗値をp値に変換する
- 6 p値から判断する

5 X二乗値をp値に変換する

これは、Excelに任せる！

関数：CHISQ.DIST.RT

X：X二乗値

自由度：1

今回のp値は 0.07364

ところで…自由度って何？！

データ値		食べた	食べてない	合計
	赤			250
	青			150
	合計	100	300	400

ここに、例えば「70」を入れると（自由に）

データ値		食べた	食べてない	合計
	赤	70		250
	青			150
	合計	100	300	400

ところで…自由度って何？！

データ値		食べた	食べてない	合計
	赤	70		250
	青			150
	合計	100	300	400

他のマスは、自動的に決まるよね？

データ値		食べた	食べてない	合計
	赤	70	180	250
	青	30	120	150
	合計	100	300	400

ところで…自由度って何？！

データ値		食べた	食べてない	合計
	赤	70	180	250
	青	30	120	150
	合計	100	300	400

自由に入れることができる場所は「1つ」

つまり、「自由度 1」

χ²乗検定の流れ

- 1 データをクロス集計表（2×2）にまとめる
- 2 期待度数を求める
（もし関係が無かったら、こうなるやろという値）
- 3 データと期待度数との差を求める（χ²乗の準備）
（この差が大きければ、関係がありそう）
- 4 χ²乗値を求める
- 5 χ²乗値をp値に変換する
- 6 p値から判断する

6 p 値から判断する

p 値

「お皿の色と食べられやすさに関係があるように、見えてしまう確率」

「有意水準」(α) 先に決めておく
普通は「5%」か「1%」

見えてしまう確率が5%

= 有意水準5%として考える

6 p 値から判断する

H_0 「お皿の色と食べるかは関係がない」

H_1 「お皿の色と食べるかは関係がある」

p 値が 5% (0.05) より小さかったら

H_0 が間違ってるとして却下できる！

つまり、

自分の言いたかった H_1 を支持する結果！

Excel シート

「お皿の色ー食べたかどうか」

「提出 検定結果」例のところ

こっちに戻って！



6 p 値から判断する

今回のp 値 **0.07364**

「色と食べられやすさに関係がある」

と見えてしまう確率 **7.3%**

5%より大きいから

H_0 が間違っているとは言えない



6 p 値から判断する

H_0 「お皿の色と食べるかは関係がない」

H_1 「お皿の色と食べるかは関係がある」

H_0 が間違っているとは言えない



「お皿の色と食べるかには、

有意な関係があるとは言えない」

×二乗検定 ポイント



データ値		食べた	食べてない	合計
	赤	70	180	250
	青	30	120	150
	合計	100	300	400

期待度数		食べた	食べてない	合計
	赤	62.5	187.5	250
	青	37.5	112.5	150
	合計	100	300	400

×二乗検定とは、

「実際のデータ値」と「期待度数（予想できる値）」が
どれだけ違ってるかを数値的に判断する！

提出 1 「性別 — 好きなご飯」

「スマホ — 甘いもの」

「睡眠時間 — 甘いもの」 の組み合わせ

表を完成させて、 χ^2 乗検定してみよう

提出 2 薬の効果 χ^2 乗検定

提出 3 近畿の男女比 χ^2 乗検定

これは、絶対練習しときや！

やってみよう