

2.ストーリー：店舗の次の目標は「ボテト売上 UP」

前回では店舗は、ボトトの重さのばらつきをデータを見て改善しました。
(本当に減ったと言ってよいか？)を判断できるようになら。

次の目標（現状の目標）

- ・ボトトは利益率が高い。
- ・だから店長はボトトの売上を伸ばすことを目標にしました。

そこで店長は考案します：

- ・「お客様を喜ぶハシゴガーナーと一緒にボトトを注文する数を増やしたい」
- ・「そのため、何が増えるとボトトも増えるのだろう？」

👉 今日ここに問い合わせて、データから予測の式を作る。

2020/01/アルゴリズム 2/10

Slide02:ストーリーの狙い（例を変えずに目的だけ変える）

【01】ボトト例を練習し、問題負荷を上げて、「回帰→予測」へ移る。

【回帰】回帰で「直線」は（重き）が重なる、今は（売上）と重なる。

【回帰】データは「回帰」：ハシゴガーナー注文数、y=ボトト注文数で記述。必ず clean で関係をつかむ。

【回帰】データは（重き）を増やすと「ボトト」が売上に反映する（増加）→「直線」。ネット注文など。

2/10

1

2

4.問い合わせ：ハンバーガーが増えるとボトトは増え？（ $x \rightarrow y$ ）

店長の目標は「ボトト売上 UP」でした。
そこで、ボトトの注文数に関心をもつてみます。

今日の問い合わせ（回帰の問題）

ハンバーガー注文数が増える日ほど、ボトト注文数も増えるのか？

回答では、2つの数値の関係を別に決めてます。

- ・x（回帰変数）：ハンバーガー注文数（前回もちもしない）
- ・y（目的変数）：ボトト注文数（増やしたい結果の側）

👉 今日やることは、xを見たらyを予測できるかを確かめること。

2020/01/アルゴリズム 2/10

Slide04:問い合わせを数学の形に落とす（混同防止）

【01】「何を予測したいか」を行って固定し、x,yの役割を先に決める。

【回帰】回帰で「直線」は（重き）が重なるを（もしない）をデータで見る。ここでは回帰は既定しない。

【回帰】clean/outlier とともに、x=ハンバーガー注文数、y=ボトト注文数で記述する（途中で変えない）。

【回帰】学生に確認：「今日のyは何？」→「ボトト注文数」。

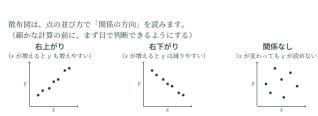
4/10

5

6

7

6.散布図の読み方①：右上がり／右下がり／関係なし（3パターン）



👉 今日のデータ（ハンバーガー→ボトト）は、右上がりになりそうかをまずは認める。

2020/01/アルゴリズム 2/10

Slide06:模式図を入れる理由（視覚で固定）

【01】散布図で「直線」を一目で判断できるように、模式図を用いて固定する。

（組みなしてみると、まずは「直線」でできる）

右上がり
(xが増えるとyも増やす)
右下がり
(xが増えるとyは減りやすい)
関係なし
(xが増えてyも増やしない)

【回帰】回帰で「直線」は（重き）が重なるを（もしない）をデータで見る。

【回帰】clean/outlier とともに、x=ハンバーガー注文数、y=ボトト注文数で記述する（途中で変えない）。

【回帰】学生に確認：「今日のyは何？」→「ボトト注文数」。

6/10

9

10

11

8.回帰の目的：点の書を「一本の線」で要約し、予測に使う

散布図を作ると、点は雲（はうらいたい集まり）になります。

このままだと、「次はどうやってよいか？」が言いたい。

回帰の目的（回帰の問題）

点の書を「一本の線」で要約して、
予測に使う形（式）にする。

👉 ここでの「線」は、全部描てる線ではなく、中心の傾向を表す線。

2020/01/アルゴリズム 2/10

Slide08:回帰が必要な理由（散布図→線への納得）

【01】なぜ線を引くのか？を言葉で説得させる（式の前の理由）。

【回帰】点がバラバラでも、中心の傾向が見えてなら直線はできる。傾きはその中心を「一本化」する。

【回帰】clean で「直線」を引くと予測できる。成績作業をする。outlier で「直線がぶる」も説明させる。

【回帰】学生に「なぜ直線の式は？」→「法線で要約して直線にする」。

8/10

13

14

10.回帰式： $y = ax + b$ （ y = 予測値）を文章で理解する

回帰直線式は、式で表せるで便利です。

（Excel が計算して表示してくれる式を、書類で説明できるのが今日の目標）

$$\hat{y} = ax + b$$

この式が何をいいこと（文句）

（ハシゴガーナー文句）がわかる。

ボトト注文数の予測値を計算できる。

（ハシゴガーナー）予測値（実測のyとは違う）

点はばらつくので、実測yは予測しないことが多い

👉 今日の実習は、Excel で式を出し、その意味を言葉で説明すること。

2020/01/アルゴリズム 2/10

Slide10:式を「暗記」させないための言い換え

【01】 $y = ax + b$ を式で見ていてなく「下の式の横をyへ」として理解させる。

【回帰】 $y = ax + b$ は「式の横をyへ」ではなく「下の式の横をyへ」として理解させる。

【回帰】近似直線の式を表すし、式に横の数を（横のy）を入れて計算させる。

【回帰】実の横の式を：「 $y = ax + b$ 」→「 $y = \text{実測} + \text{予測}$ 」。

10/10

17

18

20.回帰式： $y = ax + b$ （ y = 予測値）を文章で理解すること

回帰直線式は、式で表せるで便利です。

（Excel が計算して表示してくれる式を、書類で説明できるのが今日の目標）

$$\hat{y} = ax + b$$

この式が何をいいこと（文句）

（ハシゴガーナー文句）がわかる。

ボトト注文数の予測値を計算できる。

（ハシゴガーナー）予測値（実測のyとは違う）

点はばらつくので、実測yは予測しないことが多い

👉 今日の実習は、Excel で式を出し、その意味を言葉で説明すること。

2020/01/アルゴリズム 2/10

1.回帰で「未来予測」を作る（判断ではなく予測）

今日やること
ハシゴガーナーのデータを使って、
「ボトトの注文数を増やすには、何が関係しているの？」を見つけ、
「このくらい売れるぞ」という予測を作ります。

・今日のデータ→散布図→回帰直線→予測式 ($y = ax + b$)
・回帰は、正しい「関係」を決める話ではなく、未来を予測する話
・ゴルフ計算ではなく、式の意味を言葉で説明できること

2020/01/アルゴリズム 2/10

Slide01:第11回の立ち位置（運用）

【01】回帰（直線）から回帰（予測）へ、目的が変わることを実感する。

【回帰】回帰は「全ての点を（直線）で近づける」の直線（y=mx+b）。

【回帰】前回は clean データ（汚れを取る）で点を洗浄させる。成績は直線へ→直線。

後回し outlier データ（外れ値）で「外れ値を除く」を確認。メモ（専門用語：キーワード）と結びつける。

【回帰】測定直後に「測定直後」：今日は「直線」と「予測」のどちら？（直線）。

1/20

3.前回の検定は「判断」／今回の回帰は「予測」

前回でやっていたのは、次のようないくつか。

「平均の値がある」だけではなく、ばらつきの中で珍しいかで決めたしました。

検定（約10分）→予測
珍しい（外れ）を使って。
「何がどうあっていい／悪いか」を決める。

一方、今日次の目標です。

回帰（今日）→予測
2つの数値の関係を使って。
「もしやこのくらい、yはこのくらい」を予測する。

・検定は「重複・要除でない」で結論を出す
・回帰は「予測の式」を作り、使う（結論の用が違う）

2020/01/アルゴリズム 2/10

Slide03:混同防止（今日のモード切替）

【01】回帰（直線）と回帰（予測）を混ぜない。結果の形が違うことを固定する。

【回帰】今日の外れ値は「外れ値」ではなく、「外れ」でいいのかの心配。

【回帰】前回は「外れ」で外れ値を除く（外れを取る）で、今度は「外れ」で外れ値を除く（外れを除く）で外れを除く。

【回帰】1. 予測：検定の結果？→「直線」が直線でない。回帰の結果は？→「予測」。

3/20

5.散布図とは：2つの数値の関係を「点」で見る図（定義）

回帰を始める前に、まず散布図で目で見ます。

散布図（定義）
1つのデータ（1日）を1つ点とし、横軸x、縦軸yをとて並べた図。

今回の例では、点の意味はこうです。

・ある日：（ハンバーガー注文数x、ボトト注文数y）

・30日なら、点が30個ある

👉 散布図を見ると、増える／減る／関係ないが直感的に分かる。

2020/01/アルゴリズム 2/10

6

Slide05:散布図を先に固定する理由

【01】回帰直線の前に「点の図」を理解せない（線の意味が分からぬ）。

【回帰】回帰直線は「同時に観測させた（複数の）現象の意味が分からぬ」。

【回帰】まずは散策だけ作らせる（近似直線は次回用）。軸べらと单筋も入れせる。

【回帰】ここで小確認：「1点は何？」→「1日のデータ」。

5/20

7.散布図の読み方②：ばらつき／外れ値（現実データの特徴）

現実のデータは、きれいに一直線には見びきません。

同じxでもyが少し違うのが普通です。

重要なデータには「ばらつき」がある

ばらつきがあるなら、予測は「直線」がわけない。

でも、中の傾向をつかめば、だいたいの予測はできる。

さすが、外れ値を見つけたら「なぜ？」をメモで確認させる（説明できればOK）。

👉 後半の実習では、外れ値があると回帰直線がどう変わるかを確かめる。

2020/01/アルゴリズム 2/10

10

Slide07:ばらつきと外れ値を先に固定する（後半実習の伏線）

【01】回帰直線の前に「点の図」を理解せな（線の意味が分からぬ）。

【回帰】点がばらつく（外れ値）でない（直線）ものも、何時も「直線」があることが多い。

【回帰】前半（clean）で点を基準（直線）にする。後半（outlier）で外れ値を見せる。メモで整理づける（外れ／ノイズ／ヘビーベン）。

【回帰】外れ値を見つけたら「なぜ？」をメモで確認させる（説明できればOK）。

2020/01/アルゴリズム 2/10

11

9.回帰直線（イメージ）：各xに対して「中心」を通る予測の線

回帰では、点の中心を決めるような線を描きます。

この点で回帰直線と呼んでいます。

回帰直線（イメージ）
同じxのときの「中心」を結んだような線。

この上のyが、yの予測値（ \hat{y} ）。

・点の横の間に「ばらつき」（ばらつき）

・それで「傾向」あると、だいたいの予測がで

きる

👉 点の中心を表す線だと思えばよい（まずはこ

れでOK）。

点が散らばっていても、線があれば

「こののななうすはこのくらい」がわかる

2020/01/アルゴリズム 2/10

12

Slide09:回帰直線の直感（式数前に図で固定）

【01】回帰直線の中心の傾向（「線」上の予測値）を図で見る。

【回帰】点xでyは予測値（直線）を表す。

【回帰】手元で計算→回帰直線を表す。紙上の式が出来ることを確認させる。

【回帰】手元に書きし確認：「 $y = \text{はどこ}?$ 」→「 $y = \text{直線}$ 」。

👉 実際では、まず（増え方）を言葉で説明することを優先する。

【回帰】実際のデータがまだやまない（直線）を、紙上の式で表す。

【回帰】この直線の式を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

👉 実際では、まず（増え方）を言葉で説明することを優先する。

【回帰】実際のデータがまだやまない（直線）を、紙上の式で表す。

【回帰】この直線の式を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

2020/01/アルゴリズム 2/10

13

Slide11: aを主役、bは最小限（混乱防止）

【01】回帰直線がまだやまない（直線）を、紙上の式で表す。

【回帰】直線の式を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

【回帰】成績の傾向を（ハシゴガーナー）で起きた直感なら、成績は直線で表せる。

12. R^2 上に：線で説明できる割合の目安（正確率ではない）

判別直線を引いたら、次に気になるのはこうです。
「この割合は、どれくらいデータに合っているの？」

R²（決定係数）
回帰直線で、のばらつきのうち
どれくらいが、で説明できたかを表す。

- R^2 は 0~1 の間の数（例：0.72）
- 0に近いほど、点が線から離ればばり、説明できる割合が高い
- 0に近いほど、点が線から離ればばり、説明できる割合が低い

ここが重要な（誤解防止）
 R^2 は 正確率ではない。
「未だが必ず当たる」ことは保証しない。

Slide12 : R^2 の言い方を固定する（留学生対応）

【Q10】 R^2 を「当てた割合」か「正確率」で説明せない、直角座標上に文字で説明する。
【Q10】 R^2 が「説明できる割合」が正解でない、未だが必ず当たる保証ではない。

【Q10】 R^2 は「説明できる割合」で「 R^2 」を表示。左に R^2 と outlier で線はどう変わるか比較せよ。

【Q10】学生に「直で説明する：「 R^2 って何？」→「 R^2 で説明できる割合を日本語」」

13. 注意：相関と因果、回帰は「当てる保証」ではない

回帰ができる、つい次のように言いたくなります。
「バーガーが増えたから、カナチが増えた！」

でも注意
直角座標で回帰の式は、基本的に回帰（相関）であって、
【Q10】 R^2 を確定するどころではない。

- バーガーが増えて同時に増えた本の理由は、別にあらわせません。
- その日は来客数が多かった（回帰の誤り）
- キャンペーンをやっていた（セント注文が増える）
- 品切れがあった（方だけ減る）

今日の結果の形
回帰と言えるのか?
【Q10】 R^2 では、増える傾向がある／予測に使えるかもしれない」まで。

Slide13 : 相関と因果の線引き（やりすぎない）

【Q10】相間と因果の線引き、と誤解しない。物語で学習を教える表現に片づけられやすい。

【Q10】 R^2 が「直の正確率」で誤解される（誤解アリ）が、必ず「直角座標で説明できる」ことが必要らしい。

【Q10】ここでは説明しない。相関（因果）を含蓄として回復で持つ。

24

25

26

27

14. 実習で使うデータ：clean / outlier の2つ（目的が違う）

今日は同じデータ（オホーツク上り）でも、目的違う2種類のデータで学びます。

両者が大事です（まさか成功一次で現実）。

データA : clean（外れ値なし）
回帰の基本を「成功」させためのデータ。
散布図が読みやすく、直線の傾きがはっきり出る。

- 目的：散布図で直線一式→ R^2 を予測、スローリーズに体験
- 「傾き」の意味、言葉で説明できるようになる

データB : outlier（外れ値あり）
実際のデータで「外れ値」を体験するためのデータ。
品切れ／キャンペーンなどでが直線を外れるのが温まる。

- 目的：直線が直角座標、 R^2 と傾きに与える影響を知る
- 「外れ値」かを、メモ例から認める

2020/01/20 アホーツク上り

Slide14 : 2 データ構成の組い（成れない授業設計）

【Q10】回帰の基本を「成れ体験」で教めながら、規則（外れ値）に迷った迷れを貢献する。

【Q10】規則は同じでも、学ぶ目的違う、clean は「実習の直感」。

【Q10】outlier は「オホーツク上り」を教える（例：clean / outlier）。

【Q10】データに直線一式→ R^2 を予測できるかを確認する。

【Q10】データに直線一式→ R^2 を予測する（例：clean / outlier）

15. 実習のゴール：散布図→直線→式→ R^2 →予測→（外れ値でズレを考える）

実習では、直線を「作って使う」まで一気にやります。

最後に、外れ値で現実のズレも確認します。

- 実習の流れ（この順序で実習）**
1. 教科書を作り（ x と y の関係を点で見る）
 2. 回帰直線を引く（直角の線で一致する）
 3. 回帰式→ R^2 を表す（ R^2 →直線）
 4. R^2 を表す（除で説明できる範囲の目安）
 5. 具体的xとyで予測する（予測計算）
 6. (後半) outlier でズレの原因を考える（メモ例と結びつける）

✓ 「傾き」は直線を数値する？／「ズレた日はなぜ？」

2020/01/20 アホーツク上り

28

29

30

31

16. 実習①-1 (clean) : 散布図を作る（ x =バーガー数、 y =ボトト数）

目的：回帰はまず散布図から始める。点の並びで「関係の方向」を説む。

やること（操作）

1. clean シートを開く（外れ値なし）
2. (横軸) にバーガー注文数、 y (縦軸) にボトト注文数 選択
3. 散布図 (底) を作る

- 必ずする（第1回のルール）
 - グラフスタイル：「バーガー数とボトト数 (clean)」
 - 軸のタグ：「バーガー注文数 (件)」
 - 軸のタグ：「ボトト注文数 (件)」

✓ 「まずは左上り／右下がり／横隔なしのどれに近いかを言えるようする。」

50 / 20

Slide16 : 実習仕様（授業内の確認ポイント）

【Q10】回帰の直線は直角な直線、軸の割合を説いてみたが並んでる。

【Q10】規則は直線に上から下り（直線の傾斜）。この直線は「傾き」は言わない。

【Q10】規則は直線を引く（直角の線で一致する）。

【Q10】規則は直線を引く（直角の線で一致する）。

【Q10】直線に直角な直線を引く（直角の線で一致する）。

17. 実習①-2 (clean) : 回帰直線→式・ R^2 を表示（線形近似）

目的：点の直線一本の線で約束し、式と R^2 を表示する。

- やること（操作）**
1. 散布図の上に点をクリック
 2. 近似直線を引く（直角の線で一致する）
 3. 回帰式→式と R^2 を表示する（ R^2 →直線）
 4. 「グラフに式を表示する」をON
 5. 「グラフに R^2 値を表示する」をON

表示場所の指示「読みやすくなる」

- 式と R^2 は、直角の線で一致する（ドラッグして移動）
- 点が見えてない場合は、グラフ少し下にしてOK

✓ 「ここで見るものは2つ：回帰式（予測ルール）と R^2 （当てはまりの目安）。」

2020/01/20 アホーツク上り

32

33

34

35

18. 実習①-3 (clean) : 読み取りと予測（傾きの解釈+指定xで計算）

目的：出でた式を「読みめる」にする。

（計算よりも意味を言葉で説明できることが目標）

Step1. 傾きを「実質」で説明する
回帰式 $y = mx + b$ のは、
バーガーが1増えたとき、ボトトが平均でどれくらい増えるか。

- 例：0.6 などは、「バーガーが1増えると、ボトトは約0.6増える」

Step2. 直線で「外れ値」を計算する
外れ値を計算する（ x × $x=150$ ）を式で代入し、 y を計算する。

■■■■■の上の場所に予測（ルール）付ける。

✓ 「口頭チェック：yは実質、yは予測（同じではない）。

50 / 20

Slide18 : 授業内の確認（提出なしで確認）

【Q10】回帰の直線は直角な直線、軸の割合を説いてみたが並んでる。

【Q10】規則は直線に上から下り（直線の傾斜）。この直線は「傾き」は言わない。

【Q10】規則は直線を引く（直角の線で一致する）。

【Q10】規則は直線を引く（直角の線で一致する）。

【Q10】直線に直角な直線を引く（直角の線で一致する）。

19. 実習② (outlier) : 外れ値入りで再実行（線・式・ R^2 の変化を見る）

目的：現実のデータは「きれい」ではない。

外れ値があると、回帰直線・式・ R^2 と予測がどう変わると体験する。

- outlier シートで確認**
1. outlier シートを開く
 2. 他の回帰を引く（ターゲットも同じ）
 3. 他の回帰を引く（ターゲットも同じ）
 4. 原始の回帰直線を追加し、式と R^2 を表示する

● 説教するポイント（ここで学ぶ）

・傾きの値 = 位相は変わったか？

・傾きはどれくらい変わったか？

・ R^2 は上がった？ 下がった？

・回り x (例: $x=150$) ではどれくらい変わった？

● 理由づけ：メモ例で説明する

・外れている点を見つたら、 x と y を見て現実を言葉ににする。

■■■■■の上の場所に予測（ルール）付ける。

2020/01/20 アホーツク上り

Slide19 : 外れ値を「発見→説明」させる運用

【Q10】「外れ値」は「傾き」で説明する（直線で説明する）。

【Q10】規則は「重心の傾向」、式は「平均計算ルール」、 R^2 は「説明できる割合の目安」。

【Q10】規則は「傾き」で説明する（直線で説明する）。

【Q10】規則は「重心の傾向」で説明する（直線で説明する）。

【Q10】規則は「重心の傾向」で説明する（直線で説明する）。

【Q10】規則は「重心の傾向」で説明する（直線で説明する）。

36

37

38

39

20. まとめ

今日の副点（ここだけ）

- ・放題回：「回帰」を、回帰の基本と教える
- ・回帰式 $y = mx + b$ を使って、予測を計算できるようになる
- ・ R^2 は「当てはまりの目安」で、正確率ではない

実習のポイント（ついで）
実習で「外れ値」が「外れ値」であることを学ぶ。

次回への：では、線はどうやって決めていかれる？

Excelは線に線を引いていくよう見えるが、

実は「ズレ（残差）を小さくする」ルールで線を決めていく。

次回はその決まり→最小二乗法を学ぶ。

Slide20 : 次回（最小二乗法）への接続を自然にする

【Q10】回帰の「結果」から「決め方」を移す。

【Q10】規則は「直角の傾向」、式は「平均の傾向」 (\bar{x}, \bar{y}) 。

【Q10】規則は「重心の傾向」で説明する（直線で説明する）。

【Q10】規則は「重心の