

■ 『EXCEL 統計』『EXCEL 多変量解析』『EXCEL 官能評価』『EXCEL アンケート太閤』『Mac 多変量解析』『Mac 統計解析』の重回帰分析の動作イメージ

『EXCEL 統計 Ver. 8.0』などに搭載されている重回帰分析の動作イメージを紹介します。

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|-----|--------|-------|-----|-----|-----|---|---|---|---|---|---|---|---|
| 1 | | 選手名 | 打率 | 試合 | 本塁打 | 打点 | | | | | | | | |
| 2 | c1 | 宮崎 敏郎 | 0.326 | 124 | 20 | 71 | | | | | | | | |
| 3 | c2 | 西川 龍馬 | 0.305 | 109 | 9 | 56 | | | | | | | | |
| 4 | c3 | サンタナ | 0.3 | 136 | 18 | 66 | | | | | | | | |
| 5 | c4 | 牧 秀悟 | 0.293 | 143 | 29 | 103 | | | | | | | | |
| 6 | c5 | 大島 洋平 | 0.289 | 130 | 0 | 23 | | | | | | | | |
| 7 | c6 | 大山 悠輔 | 0.288 | 143 | 19 | 78 | | | | | | | | |
| 8 | c7 | 坂本 勇人 | 0.288 | 116 | 22 | 60 | | | | | | | | |
| 9 | c8 | 近本 光司 | 0.285 | 129 | 8 | 54 | | | | | | | | |
| 10 | c9 | 中野 拓夢 | 0.285 | 143 | 2 | 40 | | | | | | | | |
| 11 | c10 | 大城 卓三 | 0.281 | 134 | 16 | 55 | | | | | | | | |
| 12 | c11 | 岡林 勇希 | 0.279 | 143 | 3 | 31 | | | | | | | | |
| 13 | c12 | 岡本 和真 | 0.278 | 140 | 41 | 93 | | | | | | | | |
| 14 | c13 | 秋山 翔吾 | 0.274 | 115 | 4 | 38 | | | | | | | | |
| 15 | c14 | 木浪 聖也 | 0.267 | 127 | 1 | 41 | | | | | | | | |
| 16 | c15 | 坂倉 聖吾 | 0.266 | 120 | 12 | 44 | | | | | | | | |
| 17 | c16 | 佐野 恵太 | 0.264 | 141 | 13 | 65 | | | | | | | | |
| 18 | c17 | 佐藤 輝明 | 0.263 | 132 | 24 | 92 | | | | | | | | |
| 19 | c18 | 関根 大気 | 0.261 | 140 | 4 | 31 | | | | | | | | |
| 20 | c19 | 菊池 滄介 | 0.258 | 120 | 5 | 27 | | | | | | | | |
| 21 | c20 | 村上 宗隆 | 0.256 | 140 | 31 | 84 | | | | | | | | |
| 22 | c21 | 吉川 尚輝 | 0.256 | 132 | 7 | 36 | | | | | | | | |
| 23 | c22 | オスナ | 0.253 | 134 | 23 | 71 | | | | | | | | |
| 24 | c23 | 細川 成也 | 0.253 | 140 | 24 | 78 | | | | | | | | |
| 25 | c24 | 桑原 将志 | 0.252 | 132 | 7 | 35 | | | | | | | | |
| 26 | c25 | 石川 昂弥 | 0.242 | 121 | 13 | 45 | | | | | | | | |
| 27 | c26 | ノイジー | 0.24 | 133 | 9 | 56 | | | | | | | | |
| 28 | c27 | 長岡 秀樹 | 0.227 | 135 | 3 | 35 | | | | | | | | |
| 29 | p1 | 頓宮 裕真 | 0.307 | 113 | 16 | 49 | | | | | | | | |
| 30 | p2 | 近藤 健介 | 0.303 | 143 | 26 | 87 | | | | | | | | |
| 31 | p3 | 柳田 悠岐 | 0.299 | 143 | 22 | 85 | | | | | | | | |
| 32 | p4 | 森 友哉 | 0.294 | 110 | 18 | 64 | | | | | | | | |
| 33 | p5 | 松本 剛 | 0.276 | 134 | 3 | 30 | | | | | | | | |
| 34 | p6 | 紅林 弘太郎 | 0.275 | 127 | 8 | 39 | | | | | | | | |
| 35 | p7 | 中村 晃 | 0.274 | 136 | 5 | 37 | | | | | | | | |
| 36 | p8 | 浅村 栄斗 | 0.274 | 143 | 26 | 78 | | | | | | | | |
| 37 | p9 | 中川 圭太 | 0.269 | 135 | 12 | 55 | | | | | | | | |
| 38 | p10 | 万波 中正 | 0.265 | 141 | 25 | 74 | | | | | | | | |
| 39 | p11 | 辰巳 滄介 | 0.263 | 133 | 9 | 43 | | | | | | | | |
| 40 | p12 | 外崎 修汰 | 0.26 | 136 | 12 | 54 | | | | | | | | |
| 41 | p13 | マキノ | 0.259 | 127 | 15 | 50 | | | | | | | | |
| 42 | p14 | 小深田 大翔 | 0.258 | 134 | 5 | 37 | | | | | | | | |
| 43 | p15 | 今宮 健太 | 0.255 | 126 | 9 | 48 | | | | | | | | |
| 44 | p16 | マルチネス | 0.246 | 119 | 15 | 66 | | | | | | | | |
| 45 | p17 | 宗 佑磨 | 0.245 | 122 | 2 | 22 | | | | | | | | |
| 46 | p18 | ボランコ | 0.242 | 125 | 26 | 75 | | | | | | | | |
| 47 | p19 | 安田 尚憲 | 0.238 | 122 | 9 | 43 | | | | | | | | |
| 48 | p20 | 野村 佑希 | 0.236 | 125 | 13 | 43 | | | | | | | | |
| 49 | p21 | 山口 航輝 | 0.235 | 115 | 14 | 57 | | | | | | | | |
| 50 | p22 | 中村 奨吾 | 0.22 | 137 | 11 | 48 | | | | | | | | |
| 51 | | | | | | | | | | | | | | |



※説明の都合上、『EXCEL 統計 Ver. 8.0』のダイアログボックスと出力結果で説明しますが、他の製品もほとんど同じです。

※左記は 2023 年のプロ野球の打撃成績のランキングです。規定打席以上の打者が、セ・リーグが 27 人、パ・リーグが 22 人で、日本野球機構のページからコピーしました。

打撃成績の項目はいくつかあるのですが、説明の都合上、ここでは打率、試合（出場試合数）、本塁打、打点の 4 項目を使用します。

目的変数は打点、説明変数は打率と試合と本塁打の 3 項目です。

(1) メニューから「重回帰分析」を選びます。

(2) ダイアログボックスが開きますので、

- ・目的変数はセル F1～F50 の 1 列
 - ・説明変数はセル C1～E50 の 3 列
 - ・回答者ラベルはセル B2～B50 の 1 列
- をそれぞれ範囲指定します

(3) その他の指定をして「OK」ボタンを押します。

ここではデフォルトである、

- ・変数選択法 → しない
 - ・区間推定と信頼度 → なし
 - ・オプション → 散布図と正規確率プロット
- をそれぞれ選択しています。

(4) 次ページのように、計算結果が出力されます。

| | A | B | C | D | E | F | G | H |
|----|--------------|-----------|-----------|-----------|-----------|-------|--------|------|
| 1 | | | | | | | | |
| 2 | 基本統計量 | | n=49 | | | | | |
| 3 | 変数名 | 合計 | 平均 | 標準偏差 | | | | |
| 4 | 打率 | 13.122 | 0.2677959 | 0.0228117 | | | | |
| 5 | 試合 | 6398 | 130.57143 | 9.7318207 | | | | |
| 6 | 本塁打 | 668 | 13.632653 | 9.2188527 | | | | |
| 7 | 打点 | 2692 | 54.938776 | 20.134349 | | | | |
| 8 | | | | | | | | |
| 9 | 相関行列 | | | | | | | |
| 10 | | 打率 | 試合 | 本塁打 | 打点 | | | |
| 11 | 打率 | 1 | 0.019 | 0.203 | 0.280 | | | |
| 12 | 試合 | 0.019 | 1 | 0.232 | 0.312 | | | |
| 13 | 本塁打 | 0.203 | 0.232 | 1 | 0.919 | | | |
| 14 | 打点 | 0.280 | 0.312 | 0.919 | 1 | | | |
| 15 | | | | | | | | |
| 16 | 分析精度 | | | | | | | |
| 17 | 決定係数 | 0.864 | | | | | | |
| 18 | 自由度修正済み | 0.855 | | | | | | |
| 19 | ダーヴィンワトソン | 2.165 | | | | | | |
| 20 | 残差の標準偏差 | 7.673 | | | | | | |
| 21 | | | | | | | | |
| 22 | 分散分析表 | | | | | | | |
| 23 | 変動 | 偏差平方和 | 自由度 | 不偏分散 | 分散比 | p値 | 判定 | |
| 24 | 全体変動 | 19458.816 | 48 | | | | | |
| 25 | 回帰による変動 | 16809.477 | 3 | 5603.1591 | 95.171727 | 0.000 | [**] | |
| 26 | 回帰からの残差 | 2649.3389 | 45 | 58.874198 | | | | |
| 27 | | | | | | | | |
| 28 | 重回帰式 | | | | | | | |
| 29 | | 偏回帰係数 | 標準偏回帰 | F値 | p値 | 判定 | 標準誤差 | VIF |
| 30 | 打率 | 88.965 | 0.101 | 3.217 | 0.080 | [] | 49.604 | 1.04 |
| 31 | 試合 | 0.223 | 0.108 | 3.621 | 0.063 | [] | 0.117 | 1.06 |
| 32 | 本塁打 | 1.907 | 0.873 | 228.526 | 0.000 | [**] | 0.126 | 1.10 |
| 33 | 定数項 | -23.962 | | 1.420 | 0.240 | [] | 20.107 | |
| 34 | | | | | | | | |

最初に各変数の基本統計量が出力されます。2～7行目です。

次に相関係数行列が出力されます。9～14行目です。

ここまでは多変量解析の前段階である、一変量解析（単変量解析）と二変量解析で、16行目以降が重回帰分析の結果です。

16～20行目が分析精度で、

- ・決定係数（寄与率、R2 乗）
- ・自由度修正済み決定係数（自由度調整済み決定係数）
- ・ダーヴィンワトソン比
- ・残差の標準偏差

が出力されます。

18行目の自由度修正済み決定係数が0.855なので、予測精度の高い結果といえます。

22～26行目が分散分析で、セルF25のp値を確認します。p値が0.05を下回っているため、有意といえます。

28～33行目が重回帰式です。

弊社ソフトだけでなく、他社ソフトもそうだと思いますが、重回帰分析を実行しても、数式の形式では出力されません。

B列の偏回帰係数をもとに数式で表すと以下になり、これが重回帰式です。

$$y = 88.965 \times \text{打率} + 0.223 \times \text{試合} + 1.907 \times \text{本塁打} - 23.962$$

例えば、**本塁打の偏回帰係数が1.907**というのは、「**本塁打が1本増えると、打点は平均して1.907増える**」ことを意味します。

偏回帰係数は説明変数との相関係数に影響されますので、説明変数を増やしたり減らしたりした場合は偏回帰係数も変わります。

C列の標準偏回帰係数を見ると、本塁打が0.873に対して、他の2変数は約0.1でした。本塁打でほとんど説明できるといっても良さそうです。

D列のp値を見ると、本塁打は0.05を下回りましたが、打率は0.080、試合は0.063で、惜しくも有意になりませんでした。

H列のVIFを見ると、いずれも1.0少々ですので、多重共線性の心配はなさそうです。

| | A | B | C | D | E |
|----|--------|-----|----------|----------|--------|
| 35 | | | | | |
| 36 | | 実績値 | 予測値 | 残差 | 標準化残差 |
| 37 | 宮崎 敏郎 | 71 | 70.79593 | 0.204068 | 0.027 |
| 38 | 西川 龍馬 | 56 | 44.61012 | 11.38988 | 1.484 |
| 39 | サンタナ | 66 | 67.3413 | -1.3413 | -0.175 |
| 40 | 牧 秀悟 | 103 | 89.25447 | 13.74553 | 1.791 |
| 41 | 大島 洋平 | 23 | 30.70046 | -7.70046 | -1.004 |
| 42 | 大山 悠輔 | 78 | 69.73965 | 8.260355 | 1.077 |
| 43 | 坂本 勇人 | 60 | 69.44764 | -9.44764 | -1.231 |
| 44 | 近本 光司 | 54 | 45.3779 | 8.6221 | 1.124 |
| 45 | 中野 拓夢 | 40 | 37.05375 | 2.946247 | 0.384 |
| 46 | 大城 卓三 | 55 | 61.39156 | -6.39156 | -0.833 |
| 47 | 岡村 勇希 | 31 | 38.42696 | -7.42696 | -0.968 |
| 48 | 岡本 和真 | 93 | 110.1359 | -17.1359 | -2.233 |
| 49 | 秋山 翔吾 | 38 | 33.65344 | 4.346564 | 0.566 |
| 50 | 木浪 聖也 | 41 | 29.98213 | 11.01787 | 1.436 |
| 51 | 坂倉 将吾 | 44 | 49.31124 | -5.31124 | -0.692 |
| 52 | 佐野 恵太 | 65 | 55.71709 | 9.282914 | 1.210 |
| 53 | 佐藤 輝明 | 92 | 74.60079 | 17.39921 | 2.268 |
| 54 | 関根 大気 | 31 | 38.06449 | -7.06449 | -0.921 |
| 55 | 菊池 滄介 | 27 | 35.25052 | -8.25052 | -1.075 |
| 56 | 村上 宗隆 | 84 | 89.10866 | -5.10866 | -0.666 |
| 57 | 吉川 尚輝 | 36 | 41.55903 | -5.55903 | -0.724 |
| 58 | オスナ | 71 | 72.24955 | -1.24955 | -0.163 |
| 59 | 細川 成也 | 78 | 75.49277 | 2.507231 | 0.327 |
| 60 | 桑原 将志 | 35 | 41.20318 | -6.20318 | -0.808 |
| 61 | 石川 昂弥 | 45 | 49.30579 | -4.30579 | -0.561 |
| 62 | ノイジー | 56 | 44.1723 | 11.8277 | 1.541 |
| 63 | 長岡 秀樹 | 35 | 32.01917 | 2.98083 | 0.388 |
| 64 | 頓宮 裕真 | 49 | 59.02786 | -10.0279 | -1.307 |
| 65 | 近藤 健介 | 87 | 84.42311 | 2.576885 | 0.336 |
| 66 | 柳田 悠岐 | 85 | 76.43926 | 8.560744 | 1.116 |
| 67 | 森 友哉 | 64 | 61.01721 | 2.982791 | 0.389 |
| 68 | 松本 剛 | 30 | 36.15574 | -6.15574 | -0.802 |
| 69 | 紅林 弘太郎 | 39 | 44.04285 | -5.04285 | -0.657 |
| 70 | 中村 晃 | 37 | 40.23721 | -3.23721 | -0.422 |
| 71 | 浅村 栄斗 | 78 | 81.84314 | -3.84314 | -0.501 |
| 72 | 中川 圭太 | 55 | 52.91869 | 2.081313 | 0.271 |
| 73 | 万波 中正 | 74 | 78.69005 | -4.69005 | -0.611 |
| 74 | 辰巳 滄介 | 43 | 46.21849 | -3.21849 | -0.419 |
| 75 | 外崎 修汰 | 54 | 52.34071 | 1.659292 | 0.216 |
| 76 | マキノ | 50 | 55.96841 | -5.96841 | -0.778 |
| 77 | 小深田 大翔 | 37 | 38.36837 | -1.36837 | -0.178 |
| 78 | 今宮 健太 | 48 | 43.94785 | 4.052153 | 0.528 |
| 79 | マルティネス | 66 | 53.03024 | 12.96976 | 1.690 |
| 80 | 宗 佑磨 | 22 | 28.81839 | -6.81839 | -0.889 |
| 81 | ボランコ | 75 | 74.9876 | 0.012399 | 0.002 |
| 82 | 安田 尚憲 | 43 | 41.54463 | 1.455368 | 0.190 |
| 83 | 野村 佑希 | 43 | 49.66281 | -6.66281 | -0.868 |
| 84 | 山口 航輝 | 57 | 49.25381 | 7.746189 | 1.010 |
| 85 | 中村 奨吾 | 48 | 47.09782 | 0.902176 | 0.118 |
| 86 | | | | | |

最後に、各打者の**実績値** (=実際の打点) と**重回帰式から計算された予測値** (理論値) を比較した結果が出力されます。

ダイアログボックスでB列を範囲指定したことで、各サンプルがどの打者なのか識別できます。サンプル名を指定しない場合は、セルA37~A85には通し番号が出力されます。

C列の予測値というのは、前ページの数式(重回帰式)に各打者の説明変数の数値を代入した計算したものです。

D列の残差というのは、実績値と予測値との差です。

プラスであれば予測値よりも多い打点、マイナスであれば予測値よりも少ない打点であることを意味します。それぞれの絶対値が大きいほど、理論値から離れた数値といえます。

プラスの最大値(53行目)は佐藤輝明(阪神タイガース)の17.4でした。予測値が74.6打点に対して、実際には92打点をあげました。

マイナスの最大値(48行目)は岡本和真(読売ジャイアンツ)の17.1でした。予測値が110.1打点に対して、実際には93打点でした。

大切なのはE列の標準化残差で、各打者の打点に異常があるかどうかわかります。

明確な線引きはないのですが、**絶対値が3を超えていれば異常とみなすことが多い**ようです。

先ほど例に挙げた、佐藤選手と岡本選手の標準化残差はそれぞれ2.268と-2.233ですので、極端に多いまたは少ないということではないことがわかりました。

掲載は省略しますが、これらの他に、「正規確率プロット」「予測値と実測値のプロット」の2枚の散布図も出力されます。

ダイアログボックスの「区間推定と信頼度」で95%や99%を選択すると、

- (1) 偏回帰係数の95%(または99%)信頼区間
- (2) 予測値の95%(または99%)信頼区間
- (3) 個々の95%(または99%)予測区間

も出力されます。

(1)は29~33行目の標準誤差とVIFの間に、(2)と(3)はF~I列の36~85行目にそれぞれ出力されます。