『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

４．重回帰分析

4.1　重回帰分析とは

経済変数間の関数関係は，２組だけの変数だけで記述できるわけではありません．ミクロ経済で学んだように，需要を変化させる要因は価格以外に様々なものが考えられます．例えば，うどんの需要はうどんの価格以外に，所得や代替財のそばの価格や補完財のネギの価格などの需要を変化させる要因があります．

このように，説明変数が２つ以上の場合を**重回帰分析**と言います．基本的な考え方は単回帰分析と同じです．i 番目に観測された変数 y_i（例えば，うどんの需要量）が k 個の変数 x_{i1}（例えばうどん価格），x_{i2}（所得）x_{i3}（そばの価格），･･･，x_{ik} と関係しており，

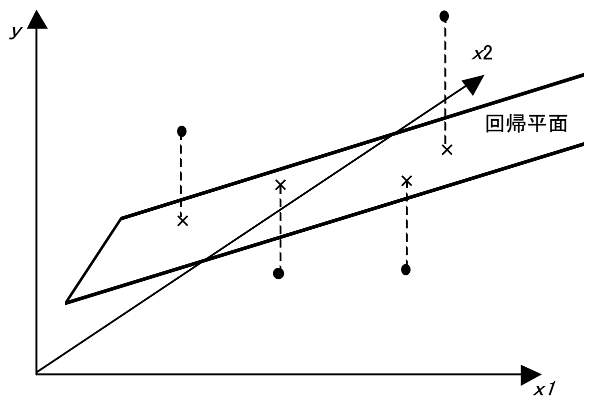
線型モデル：$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}$　(1a)

のように，定数（β_0）と $x_{i1}, x_{i2}, x_{i3},\cdots, x_{ik}$ の線形式で表されるとします．しかし，(1a)式の線型モデルには，定量的に把握することができないデータ（おなかの減り具合など）の影響が考慮されていません．そこで，i 番目の観測データ（$x_{i1}, x_{i2}, x_{i3},\cdots, x_{ik}$）では把握できないバラツキである ε_i を考慮したモデルが，次式の確率モデルです．

重回帰式：$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i$　(1b)

以下に説明する方法によって，観測データをモデルにあてはめて，変数 y_i を説明変数 $x_{i1}, x_{i2}, x_{i3},\cdots, x_{ik}$ によって説明しようとすることを「y を定数と $x_{i1}, x_{i2}, x_{i3},\cdots, x_{ik}$ に重回帰する」と言い，この式を**重回帰式(重回帰方程式)**と言います．**重回帰式**で β_0 を定数項，$\beta_1,\cdots,\beta_k$ を**回帰係数**（**偏回帰係数，母回帰係数**）と言います．

図 4-1　　重回帰方程式と残差



4.2　重回帰式の推計（大学院に行こうという学生はレジュメ最後のページも読んでください．）

単回帰と同様に，

$$s(\beta_0, \beta_1, \cdots, \beta_k) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_{1i} - \cdots - \beta_k x_{ki} \right)^2 \quad (2)$$

の残差平方和を最小にする $\beta_0, \beta_1,\cdots,\beta_k$ を求めます（単回帰と同様に**最小二乗法**といいます）．ようするに，最小二乗法から求められる回帰係数は，上図で，回帰平面で求められる**理論値(あてはまり値)**と**観測値**の距離（残差）の合計を最短にします．詳しい考え方はレジュメ最後のページを読んでください．単回帰のように簡単な式で解を示すことは出来ませんが，重回帰式も単回帰式同様に，以下のように各変数のデータの平均値を通る特性があります．

$$\bar{y} = b_0 + \beta_1 \bar{x}_1 + \cdots + \beta_k \bar{x}_k \quad (3)$$

4.3　重回帰式のあてはまり

重回帰式のあてはまりも単回帰と同じような考えです．

$$\hat{y}_i = b_0 + b_1 x_{1i} + \cdots + b_k x_{ki} \quad (4)$$

$$e_i = y_i - \hat{y}_i \quad (5)$$

『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

重回帰式の特性
① 平均 $\left(\bar{y}, \bar{x}_1, \cdots \bar{x}_k\right)$ を通る.
② 残差の総和，平均は０.
③ 残差と説明変数は直交する

重回帰式でも，下式は成立します.

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+\sum_{i=1}^{n}e_i{}^2 \tag{6}$$

全変動＝回帰変動＋残差変動

決定係数$=\mathrm{R}^2=$(*回帰変動*)/(*全変動*)

しかし，ここでこのまま**決定係数**を用いると問題があります．残差変動を*RSS*(*k*)とすると

$$R^2(k)=1-\frac{RSS(k)}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{7}$$

となります．ここで，説明変数を１個追加すると，残差変動は*RSS*(*k*+1)となり，*RSS*(*k*+1)＝min S（β_0, β_1, ････, β_h, β_{h+1}）≦min S（β_0, β_1, ････, β_h, ０）＝*RSS*(*k*)から $R^2(k)\leq R^2(k+1)$ となります．**ようするに，これは被説明変数と何の関係もない変数を説明変数に追加しても決定係数が上がることになります(ここだけ理解してください).** そこで，

$$\bar{R}^2=1-\frac{\sum_{i=1}^{n}e_i^2\Big/(n-k-1)}{\sum_{i=1}^{n}(y_i-\bar{y})^2\Big/(n-1)} \tag{8}$$

で表される，**説明変数の個数が与える影響を考慮した決定係数である自由度修正済決定係数($\bar{R}^2$)を使います.** ここで，ｋは説明変数の数．自由度修正とは決定係数の定義において残差変動と全変動の比をそれぞれの自由度（n-k-1)，(n-1)で割った平均変動の比に置き換えたもの．残差変動は必ずしも減少せず，全変動の値は説明変数の追加，除去と無関係．説明変数の追加，除去の判断には $\bar{R}^2$ を使います.

4.4　Excelによる重回帰分析

さて，重回帰分析は単回帰分析のように表から簡単に計算できるものではありません．以後は，Excelのアドインソフトの一つの分析ツールを用いて重回帰分析を行っていきます.

表4-1は新浦安駅周辺のコンビニ各店の売り上げと店舗の特性です．１日あたりの売り上げ（ｙ）は店舗面積と駅からの距離と関係があるかを**重回帰分析**で調べてみましょう.

表 4-1　新浦安駅周辺のコンビニの売り上げ

店舗	一日平均売上高（万円） y	店舗面積（㎡） x１	駅からの距離（100m） x２
A	４０	６０	3
B	４５	１００	5
C	８０	８５	2
D	６０	５０	1
E	５０	７５	3
F	２０	５５	4
G	１５	７０	6
H	９０	９５	1
I	３０	４５	3
J	７０	６５	2

まず，Econome5.xlsをＤＬして，****data05.xlsと名前を変えて各自のＵＦ（ユーザーフォルダー）の計量分析の中にセーブしてください．****data05.xlsのsheet name　“コンビニの売り上げ”をアクティブにしてください．Excelのツールバーから分析ツールを選択し，メニューリストから回帰分析を選んでください

作業 1.．****data05.xls の sheet name　“コンビニの売り上げ”を用いてコンビニの売り上げを，単回帰モデル（説明変数を①店舗面積のみで行った場合と②駅からの距離で行った場合の２種類）と重回帰モデル（説明変数として，店舗面積と駅からの距離の２つを使う場合）を最小二乗法で推定してください．全部で３つありますので，結果を比較できるように横に並べて出力してください．詳しい操作の仕方は，下図をみよう．（５点×３）

『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

図 4-2　Excel 分析ツール

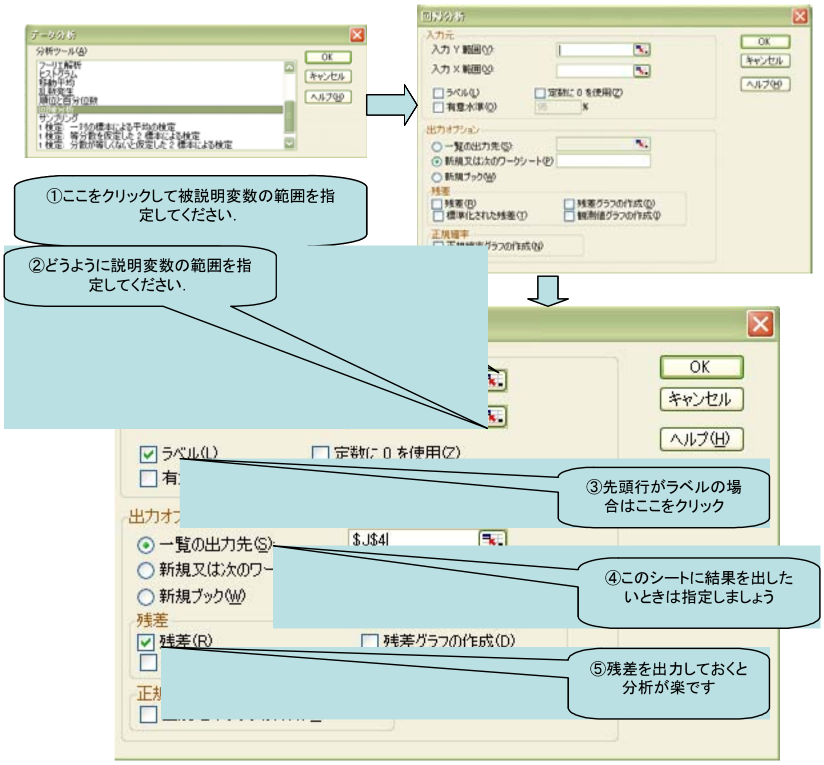


図 4-3　Excel の重回帰分析結果

店舗	一日平均売上高 万円 y	店舗面積 ㎡ x1	駅からの距離 100m x2
A	40	60	3
B	45	100	5
C	80	85	2
D	60	50	1
E	50	75	3
F	20	55	4
G	15	70	8
H	90	95	1
I	30	45	3
J	70	65	2

概要

回帰統計	
重相関 R	0.485931
重決定 R2	0.236129
補正 R2	0.140645
標準誤差	23.2268
観測数	10

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	1	1334.127	1334.127	2.472968	0.154464
残差	8	4315.873	539.4841		
合計	9	5650			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	4.444444	29.88555	0.148716	0.885458	-64.4718	73.36064	-64.4718	73.36064
x1	0.650794	0.413842	1.572567	0.154464	-0.30353	1.605114	-0.30353	1.605114

残差出力

観測値	予測値 : y	残差
1	43.49206	-3.49206
2	69.52381	-24.5238
3	59.7619	20.2381
4	36.98413	23.01587
5	53.25397	-3.25397
6	40.2381	-20.2381
7	50	-35
8	66.26984	23.73016
9	33.73016	-3.73016
10	46.74603	23.25397

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

4.5　回帰分析の意味（**ここは重要！！　今後の提出物には，以下で説明する「回帰分析の出力結果の記述法」「回帰係数の意味（解釈）」を必ず書くこと．なければ0点ですよ．．．**）

Excel の回帰分析ツールで出力される結果の読み方について確認しましょう．

回帰分析では，変数間にどんな関係があるかデータをもとに調べることができます．作業１では，「コンビニの一日平均売上高」と「店舗面積」および「駅からの距離」の関係を調べるために回帰式を推定しました．コンビニの立地計画をたてるには，「店舗面積」や「駅からの距離」が変化すると「コンビニの一日平均売上高」がどれだけ変わるか知る必要があります．ここで，「店舗面積」や「駅からの距離」を**説明変数**と言います．また，説明変数の変化による影響をみたい「コンビニの一日平均売上高」を**被説明変数**と呼びます．説明変数が１つだけの回帰分析を単回帰分析，説明変数が２つ以上ある回帰分析を重回帰分析と呼びます（注：ここでは定数項は数に入れていません）．以下，「コンビニの一日平均売上高」と「店舗面積」および「駅からの距離」の関係をみる重回帰分析を例としてみていきましょう．

それでは，Excel の回帰分析ツールの流れに沿って確認しましょう．まず，入力画面（「ツール」－「分析ツール」－「回帰分析」）に注目してください．次の２つの項目があります．

入力Ｙ範囲
入力Ｘ範囲

入力Ｘ範囲には，説明変数のデータを選択します．入力Ｙ範囲には，被説明変数のデータを選択します．作業１の重回帰モデルの場合には，入力Ｘ範囲には，「店舗面積」と「駅からの距離」のデータを選び，入力Ｙ範囲には，「コンビニの一日平均売上高」のデータを選びました．なお，入力画面の「ラベル」がチェックされている場合には，Excel は，データの最初の行をデータではなく変数名として扱いますので，注意してください．入力画面を正しく設定すると，次の出力が出ます．

図 4-4　回帰分析の出力例

概要

回帰統計	
重相関 R	0.9780619
重決定 R2	0.956605
補正 R2	0.9442064
標準誤差	5.9182735
観測数	10

重回帰式のあてはまりを検討する

分散分析表

	自由度	変動	分散	観測された分散	有意 F
回帰	2	5404.818	2702.4091	77.15446	1.7E-05
残差	7	245.1817	35.025961		
合計	9	5650			

各説明変数の説明力を検討する

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	36.412	8.172	4.456	0.003	17.089	55.736	17.089	55.736
x1	① 0.755	0.106	③ 7.126	0.000	0.504	1.005	0.504	1.005
x2	② -13.078	1.213	④ -10.781	0.000	-15.946	-10.209	-15.946	-10.209

残差出力

観測値	予測値 : y	残差
1	42.454151	-2.45415
2	46.482161	-1.48216
3	74.396465	5.603535
4	61.063688	-1.06369
5	53.772924	-3.77292
6	25.603535	-5.60353
7	10.766922	4.233078
8	95.020007	-5.02001
9	31.135378	-1.13538
10	59.304768	10.69523

異常値があるか検討する

出力結果の読み方

（１）推計式のあてはまり

まず，点線１で囲った表をみてください．概要表はこの重回帰分析のあてはまりを検討しているところです．ここで，重要になるのは補正 R2 と記してある，**自由度修正済み決定係数** $\overline{R}^2$ です．これはこの重回帰式の“あてはまり”を表していました．**自由度修正済み決定係数** $\overline{R}^2$ が 0.9442064 と１に近いですから，“あてはまり”は高いと言えます．

『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

（２）説明変数の読み方

次に，波線２で囲った表を見てください（この出力は，見やすさを考えて小数点以下３桁まで表示させています）．この表は，次のように読みます．

回帰係数の意味（解釈）

まず，X1 の行は，入力画面で入れた最初の（いちばん左側の列の）説明変数についての結果を示します．この課題の場合は，「店舗面積」でした．次に，X2 の行は，２番目の説明変数について結果で，これは，「駅からの距離」を表しています．

表の中の「係数」に注目しましょう．これは，授業で習った**回帰係数**のことです．**この意味を読み取りましょう**．**まず，X1 の係数（①のセル 0.755）は，他の条件を一定とした場合に「店舗面積」を１単位増やすと「コンビニの一日平均売上高」が何単位変化するか**を示しています．ここでは，「コンビニの一日平均売上高」に影響を与える条件（説明変数）として「店舗面積」と「駅からの距離」だけを考えていますから，①のセルの係数は，「駅からの距離」を同じにして「店舗面積」だけを１単位増やした場合に，「コンビニの一日平均売上高」が何単位変化するかを示します．ここで１単位とは，計算に使ったデータの単位を示します．「コンビニの一日平均売上高の 1 単位は「1 万円」，店舗面積の 1 単位は「１㎡」でした．回帰係数の数字は，0.755 ですから，「店舗面積」が１㎡増えれば，「コンビニの一日平均売上高」は，0.755 万円増えると読み取れます．

同じように，「店舗面積」を一定として「駅からの距離」だけを増やした場合の「コンビニの一日平均売上高」の変化は，X2 の係数（②のセル）から読み取れます．この数字は，－13.078 ですから，「駅からの距離」を１単位（100m）伸ばすと，売上高が 13 万 780 円減ることが分かります．このようにして回帰係数から読み取りますので，データの単位やオーダーはきちんと把握しておきましょう．

回帰係数の統計的意味

次に，上で見た関係が，統計的に意味があるか調べてみましょう．まず，「店舗面積」です．「店舗面積」が１㎡増えれば，「コンビニの一日平均売上高」は，0.755 万円増えるという結果が出ましたが，これは，たまたまこのデータだけでみられた偶然の結果かも知れません．売上高のデータをとりなおせば，また違う結果が出るかも知れません．そこで，「店舗面積」と「コンビニの一日平均売上高」の間に本当に関係があるのか，統計的なチェックが必要なのです．これは，X1 の t 値を見ることでチェックできます．この値の絶対値が２をこえていれば，「店舗面積」と「コンビニの一日平均売上高」の間の関係が統計的に確認できたと考えます．この場合，「店舗面積」の t 値は 7.126（③のセル）ですから，余裕で２をこえています．このため，**「店舗面積」と「コンビニの一日平均売上高」には統計的に意味のある関係**があるといってよいでしょう．

同様に，「駅からの距離」と「コンビニの一日平均売上高」の関係もチェックします．こちらの t 値は，-10.781（④のセル）です．絶対値は 10.781 ですから，やはり，余裕で２をこえています．このため，「駅からの距離」と「コンビニの一日平均売上高」の間にも統計的に意味のある関係があると言えます．**一般に，ｔ値の絶対値が２以上であれば回帰係数は有意である（その回帰係数は０ではない，すなわち，回帰式にその説明変数が必要である）と判断されます．**

この出力表に，下限 95%，上限 95%とありますが，これも参考になる指標です．例えば，「店舗面積」Ｘ１の回帰係数は 0.755 で，下限 95%が 0.504，上限 95%が 1.005 となっています．これは，「店舗面積」Ｘ１の回帰係数が 95%の確率で 0.504 から 1.005 の間にあることを示します．この範囲がゼロを含まないことから，回帰係数はゼロではない可能性が高く，これからも説明力の高さが分かります．

（３）観測データ別理論値と残差

最後に，残差出力表（一点波線で囲ってある部分）があります．これは，データ別に予測値（**理論値**）と**残差**が示されています．観測値と書いてあるのはここでは店舗番号です．基データの順番で店舗Ａが１，店舗Ｂが２となっています．例えば，店舗Ａでは，観測値が 40 万円で，理論値 42.45 万円ですから**残差**（**理論値－観測値**）が-2.45 万円となります．この出力表では，個々の観測データを考慮して，追加すべき説明変数を検討したり，個々のデータの**異常性**を検討します．追って，これらを説明しますが，例えば観測値が 10 となっている店舗Ｉでは残差が大きくなっていますので，その原因を探る必要があるというような検討です．

作業 2. rep05a.doc をＤＬして，****rep05a.doc と名前を変えて各自のＵＦの計量分析の中にセーブしてください．****rep05a.doc に書いてある課題は，作業１の３つの回帰式の推定結果を使って答えてください．****は学籍番号．

『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

さて，作業２までの分析で行ったことを式で確認しておきましょう．これまでの講義の学んだことがすべて出てきますので，わからないことはここでクリアにしておこう．作業１の重回帰分析では，次の回帰式を想定しています．

(コンビニの一日平均売上高) ＝
　　(切片) ＋ (店舗面積の回帰係数)×(店舗面積) ＋ (駅からの距離の回帰係数)×(駅からの距離) ＋ 誤差項（ε）

Excel は，入力したデータをもとに，上の式の(切片)，(店舗面積の回帰係数)，(駅からの距離の回帰係数) を**最小二乗法**という手法で計算し，出力してくれているのです．Excel が求めた値を代入すると，

(コンビニの一日平均売上高) ＝
　　36.412 ＋ 0.755×(店舗面積) － 13.078×(駅からの距離の回帰係数)×(駅からの距離) ＋ 誤差項（ε）

となります．

回帰分析の出力結果の記述法

卒業論文などで回帰式の推計結果を報告する場合には、次のいずれかの形式にしてください．

◆**式形式**：式形式の場合には，推定した回帰係数を使って回帰式を書く．また，回帰係数の下には括弧書きで t 値を書き，回帰式の下に自由度修正済み決定係数，標準誤差，観測数などを書きます．式形式のメリットは，回帰式の構造を把握しやすいことであり，デメリットは，説明変数が多い場合や複数タイプのモデルを同時に書きにくいことです．

(コンビニの一日平均売上高（万円））＝36.412＋0.755×(店舗面積)－13.078×(駅からの距離)
　　　　　　　　　　　　　　　　　(4.46)　　(7.13)　　　　　(-10.78)

$\hat{R}^2$＝0.944　標準誤差＝5.918　観測数＝10

◆**表形式**：表形式の場合には，下表にあるように回帰係数ごとに推定値， t 値（，標準誤差）を書きます．また，その下などに自由度修正済み決定係数，標準誤差，観測数などを書きます．表形式のメリットは，説明変数が多い場合や複数タイプのモデルを同時に記述できることであり，デメリットは式形式のように回帰式の構造を把握しやすくないということです．

表 4-2　コンビニの売り上げ推計結果（万円）

	モデル1		
	回帰係数	標準誤差	t 値
定数項	36.412	8.172	4.46
店舗面積	0.755	0.106	7.13
駅からの距離	-13.078	1.213	-10.78
自由度修正済み決定係数		0.944	
標準誤差		5.918	
観測数		10	

説明変数は縦に書き足していきます.

複数のモデルは横に書き足します.

回帰分析の応用

ここで，店舗面積 80 平方メートルのK店を駅から 300mの地点に出店したときの一日平均売上高を予測してみましょう．やり方は簡単で，店舗面積と駅からの距離（100m 単位）を前のページの最後の式に入れるだけです．すると，コンビニ一日平均売上高は，

　　36.412 ＋ 0.755×80 － 13.078×3 ＋ 誤差項（ε）

となることがわかります．問題は誤差項でこの大きさはわからないのですが，前回までのレジュメにある条件がみたされていれば，このお店の一日平均売上高の予測では，誤差項を無視してかまいません．このため，予測結果は，以下の値となります．（正確には，誤差項を無視して良いのは，一日平均売上高の平均値を予測する場合です．たとえば，一日平均売上高が毎月計算される値だとすると，その 5 年分の平均値を予測する場合です．ただし，このような込み入った話はこの授業の段階では考えないことにしましょう）

　　36.412 ＋ 0.755×80 － 13.078×3＝57.6（万円）

（やや上級の話題）
上で用いた回帰式は，説明のわかりやすさを考えてもっとも簡単な形にしてあります．このため，やや非現実的な特徴があります．

『不動産のための計量分析』レジュメ NO.5

クラス担当教員名※1	学部・学科名	学籍番号※2	氏　名※2	提出日

※1:履修登録したクラスの担当教員名を書く　※2:学籍番号及び氏名が未記入のもの，また授業終了後に提出されたものは採点しないので，注意すること.

まず，もともとの広さが 10 ㎡の店を 1 ㎡広げる場合も，200 ㎡の店を 1 ㎡広げる場合も，売り上げの増え方は同じ 0.755 万円と計算されています．また，駅前（駅からゼロ m）の店を 1 ㎡広げる場合も，駅から 1 km の店を 1 ㎡広げ場合も売り上げの増え方はやはり同じ 0.755 万円です．これは，ちょっと不自然な気もしますね．説明変数や被説明変数に変数変換をほどこしたり，回帰式の形を変えたりすることで，もっと現実的な結果を出すことができるようになります．これについては，次章で説明していきます.

作業 3. rep05b.doc をＤＬして，****rep05b.doc と名前を変えて各自のＵＦの計量分析の中にセーブしてください．****rep05b.doc にある応用問題は，作業１の重回帰分析の結果を使って答えてください．****は学籍番号.

【参考:大学院への進学を考えている人は読んでね】

単回帰と同様に，残差平方和

$$s(\beta_0,\beta_1,\cdots,\beta_k)=\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_{1i}-\cdots-\beta_kx_{ki}\right)^2 \qquad (2)$$

を最小にする β_0，β_1，・・・・，β_k を求めます（で，単回帰と同様に**最小二乗法**といいます）．ようするに，最小二乗法から求められる回帰係数は，上図で，回帰平面で求められる**理論値(あてはまり値)**と**観測値**の距離（残差）の合計を最短にします.

定数項と回帰係数は $s(\beta_0,\beta_1,\cdots,\beta_k)$ を β_0，β_1，・・・・，β_k で偏微分し，ゼロとすると，ｋ＋１本の式（正規方程式）から求めます.

$$b_0\sum_{i=1}^{n}1+b_1\sum_{i=1}^{n}x_{1i}+\cdots+b_k\sum_{i=1}^{n}x_{ki}=\sum_{i=1}^{n}y_i$$

$$b_0\sum_{i=1}^{n}x_{1i}+b_1\sum_{i=1}^{n}x_{1i}^2+\cdots+b_k\sum_{i=1}^{n}x_{1i}x_{ki}=\sum_{i=1}^{n}x_{1i}y_i$$

$$\cdots\cdots$$

$$b_0\sum_{i=1}^{n}x_{ki}+b_1\sum_{i=1}^{n}x_{1i}x_{ki}+\cdots+b_k\sum_{i=1}^{n}x_{ki}^2=\sum_{i=1}^{n}x_{1ki}y_i$$

この正規方程式はｋ＋１個の未知数でｋ＋１本の連立方程式であるので解が得られる条件は説明変数ｘ1，ｘ2，･･，ｘｋと定数１に一次従属の関係がないことです．一次従属の場合は多重共線の関係と言います．単回帰のように簡単な式で解を示すことは出来ませんが，最初の式をｎで割ると，以下のように各変数のデータの平均を通ることが分かります.

$$\overline{y}=b_0+\beta_1\overline{x}_1+\cdots+\beta_k\overline{x}_k \qquad (3)$$