

回帰分析(Linear Regression)

xの値が決まるとyの値が決まる、xが変化するとyも変化する
=xはyの独立変数（説明変数）、yはxの従属変数（被説明変数）

独立変数xと従属変数yの間の関係を数式を用いて表現すると、

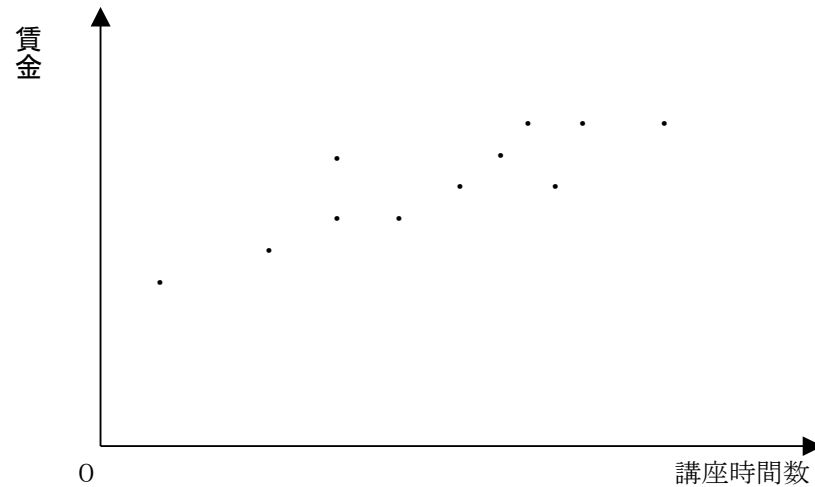
$$y = a + bx \quad (a, b: \text{定数})$$

回帰分析とは、政策（施策、事業）が行われることでxの値が変化し、その結果yの値が変化するという、

原因（独立変数xの変化）→結果（従属変数yの変化）
に至る両者の間の因果関係を見つけだし、これを方程式体系で表現すること

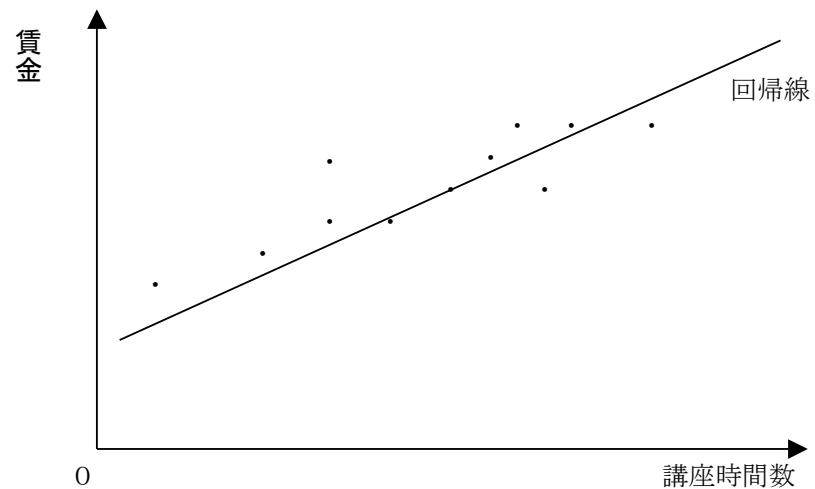
独立変数としての政策（施策、事業）の実施と従属変数としての目的変数の間における因果関係の存在を検証するための分析方法

就職支援講座受講時間数と受講後の賃金の関係



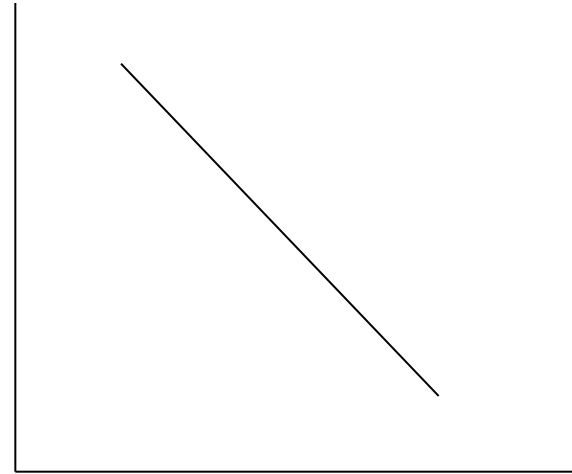
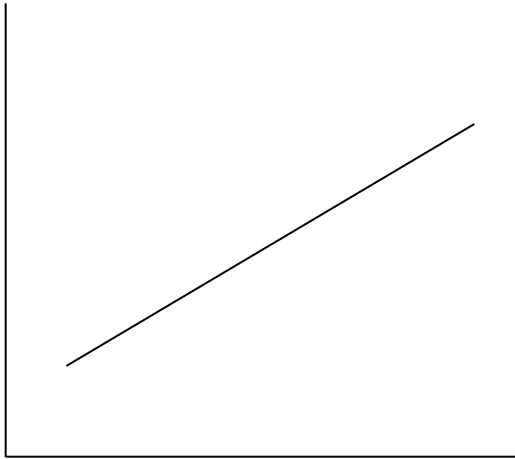
受講時間数→賃金
(x) → (y)

傾向線（回帰線）を
見つけ出す

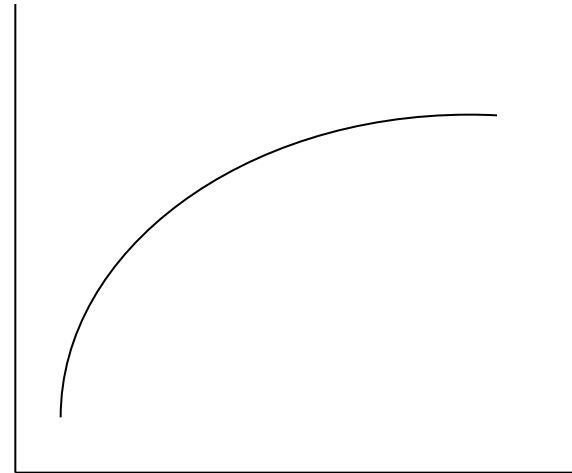
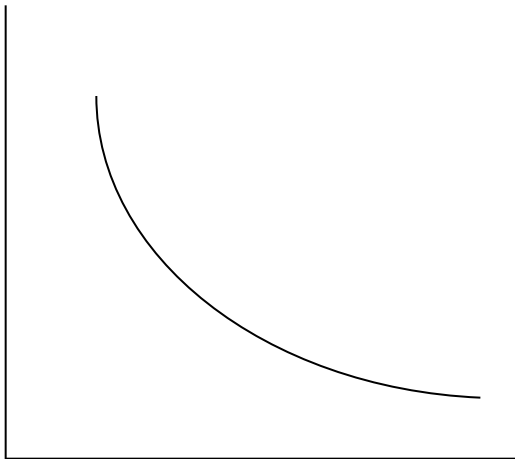


$$y = 10 + 0.8x$$

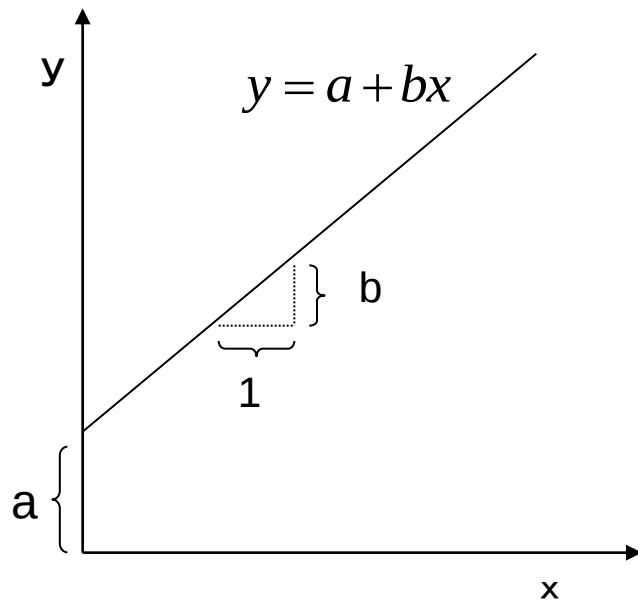
線形回帰 (Linear Regression)



非線形回帰 (Non-Linear Regression)



回帰方程式の算出



係数 : b

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

切片(定数項) : a

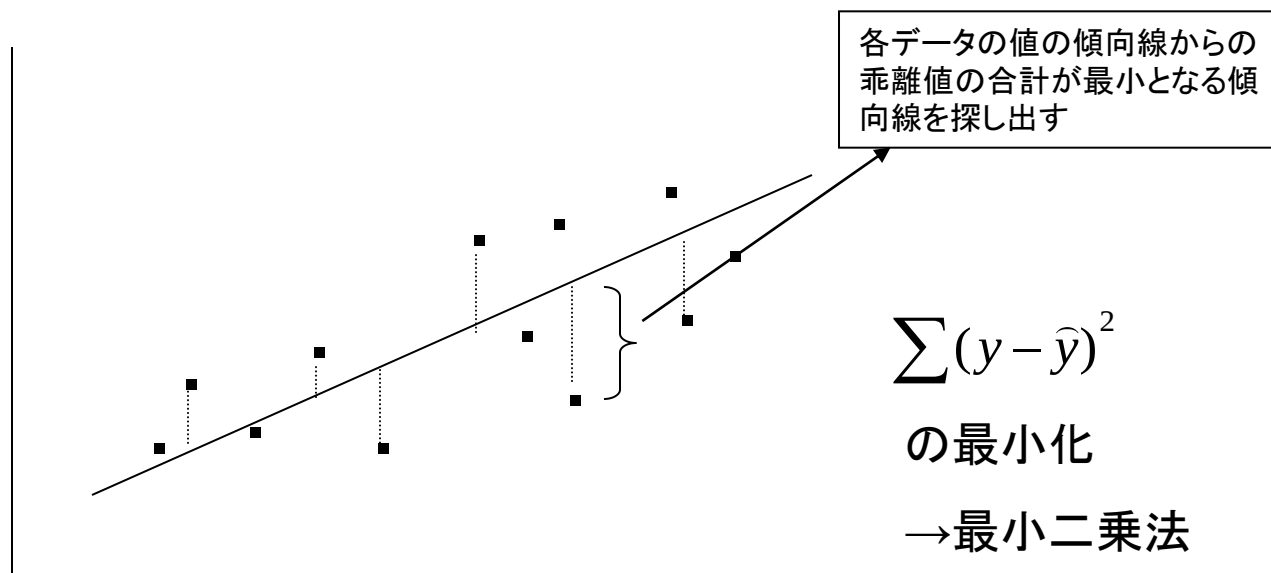
$$a = \bar{y} - b\bar{x}$$

講座時間数 x	0 時間	1 時間	5 時間	10 時間	20 時間	30 時間
賃金 y	800 円	810 円	850 円	900 円	1,000 円	1,100 円

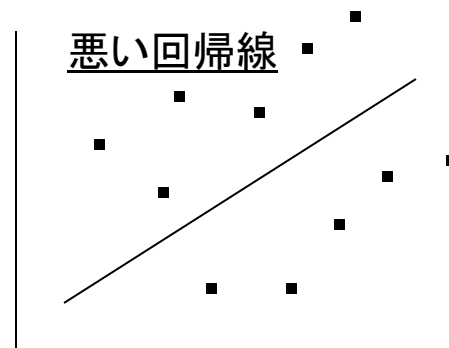
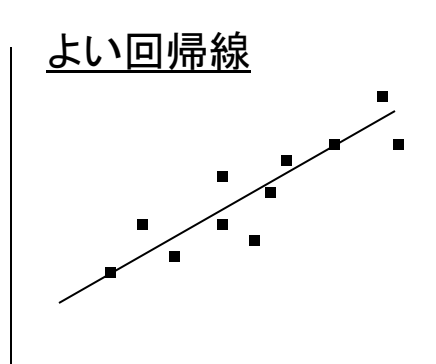
(
x が 1 単位増えると、
y は 10 (1 × b) 増える

(
x が 10 単位増えると、
y は 100 (10 × b) 増える

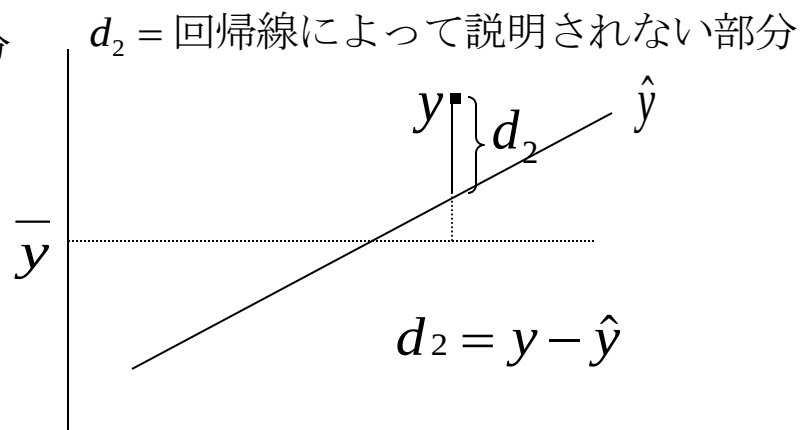
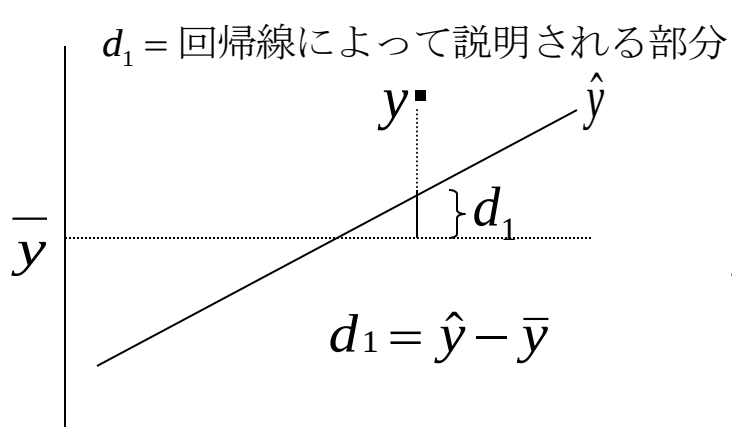
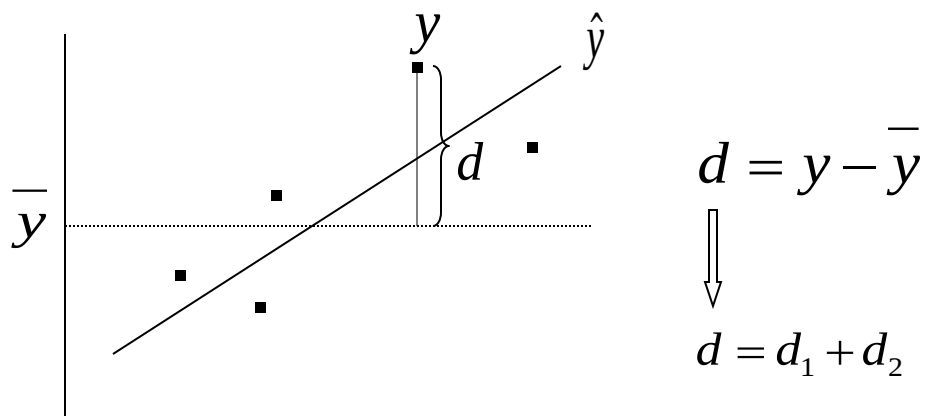
① 最小二乗法—どのようにして回帰線を見つけるか



② 当てはまりのよい回帰線と当てはまりの悪い回帰線

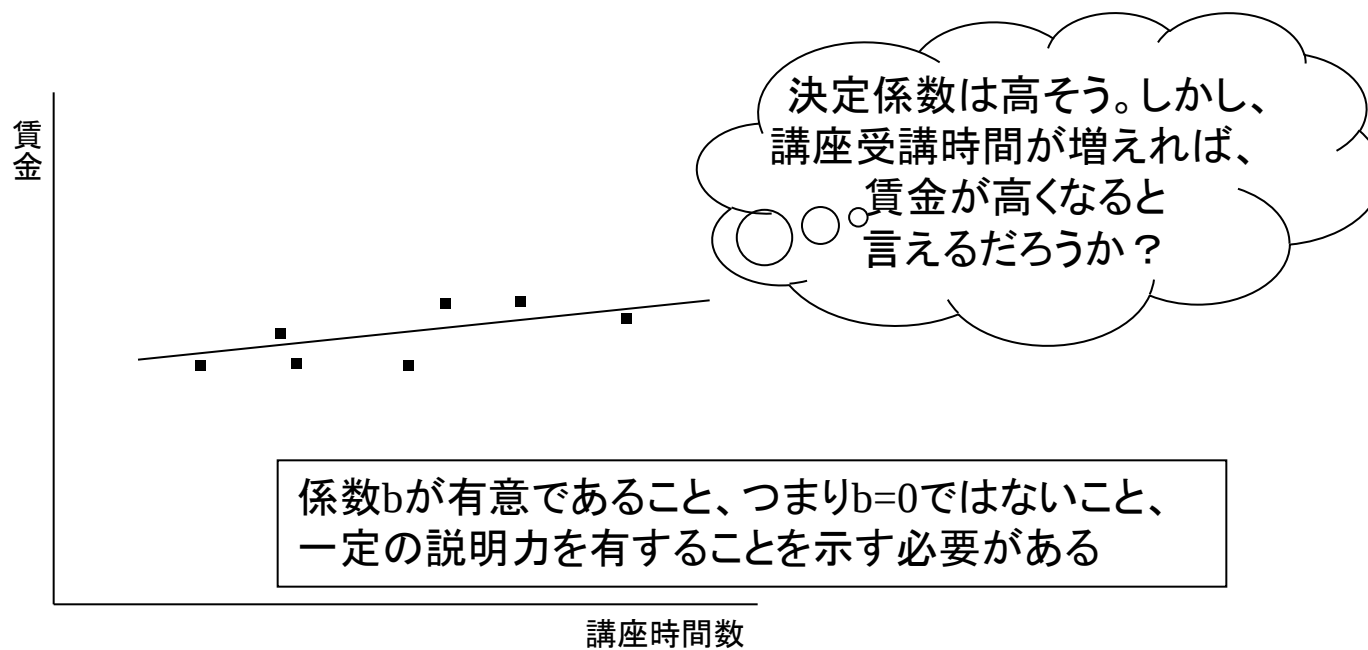


③ 決定係数



$$\text{決定係数 } R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}$$

④ t値



t値 = 係数の標準得点

⑤ SPSSを使った単回帰分析

講習参加時間と作業ミスの件数の関係

ID	参加時間 (時間)	件数 (回／年)
A	30	80
B	25	80
C	50	45
D	38	70
E	20	95
F	70	20
G	35	50
H	24	90
I	60	25
J	45	50

*無題1 [データセット0] - IBM SPSS Statistics データ エディタ

ファイル(F) 編集(E) 表示(V) データ(D) 変換(T) 分析(A) ダイレクト マーケティング(M) グラフ(G) ユーティリティ(U) ウィンドウ(W) ヘルプ(H)

表示: 2 個 (2 変数中)

	時間	件数	var
1	30	80	
2	25	80	
3	50	45	
4	38	70	
5	20	95	
6	70	20	
7	35	50	
8	24	90	
9	60	25	
10	45	50	
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			

分析(A) のメニュー:

- 報告書(P)
- 記述統計(E)
- テーブル(B)
- 平均の比較(M)
- 一般線型モデル(G)
- 一般化線型モデル(Z)
- 混合モデル(X)
- 相関(C)
- 回帰(R)**
 - 自動線型モデリング...
 - 線型(L)...**
 - 曲線推定(C)...
 - 偏相関最小 2 乗法(S)...
 - 二項ロジスティック(G)...
- 対数線型(O)
- 分類(F)
- 次元分解
- 尺度(A)

線型回帰

従属変数(D): 件数

ブロック 1 / 1

独立変数(I): 時間

方法(M): 強制投入法

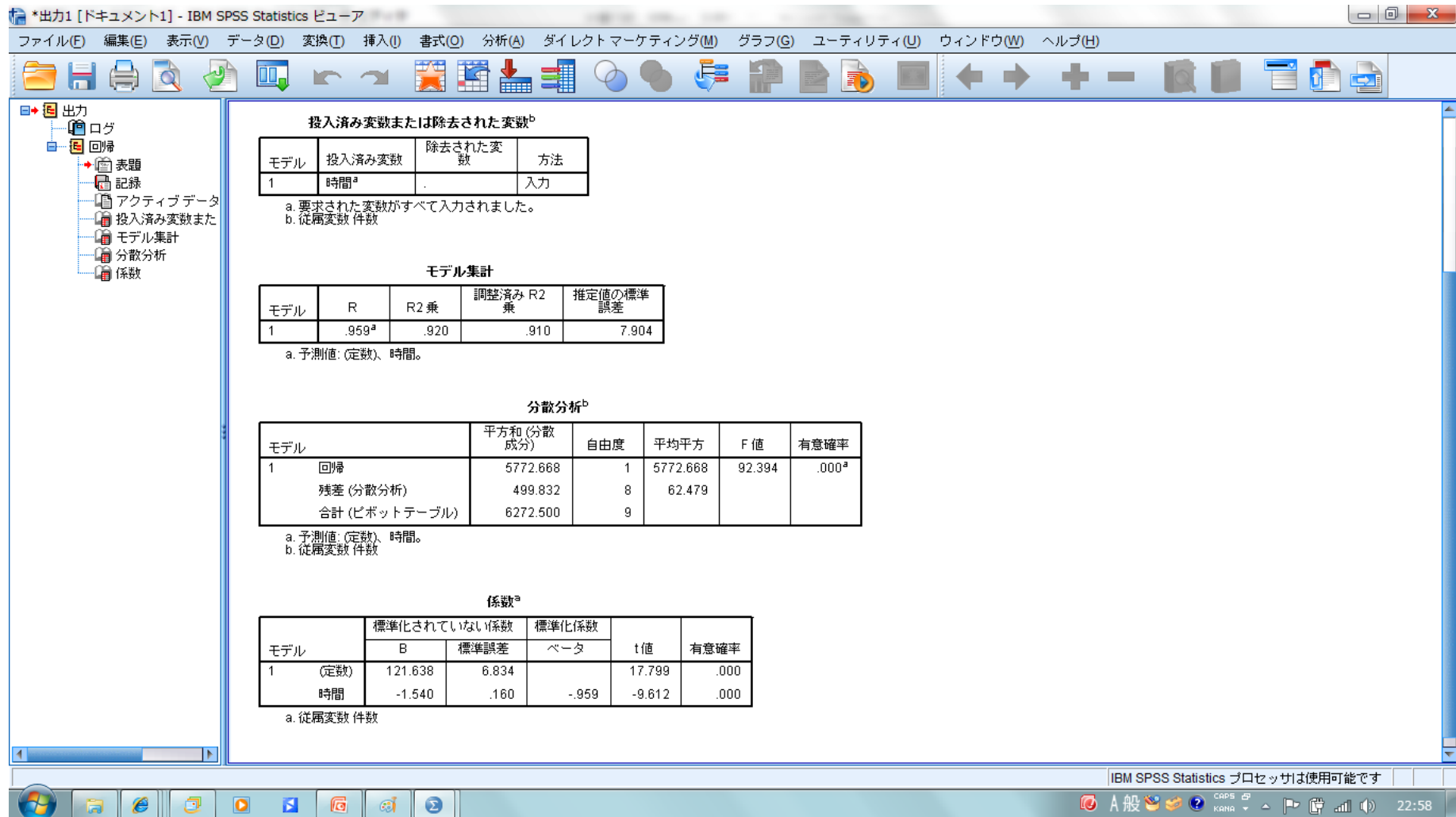
ケース選択変数(E):

ケースのラベル(C):

WLS 重み(H):

統計量(S)...
作図(T)...
保存(S)...
オプション(O)...

OK 貼り付け(P) 戻す(R) キャンセル ヘルプ



*出力1 [ドキュメント1] - IBM SPSS Statistics ビューア

ファイル(F) 編集(E) 表示(V) データ(D) 変換(T) 挿入(I) 書式(O) 分析(A) ダイレクトマーケティング(M) グラフ(G) ユーティリティ(U) ウィンドウ(W) ヘルプ(H)

図表ビルダー(C)...
 グラフボード テンプレート選択(G)...
 レガシー ダイアログ(L)

棒(B)...
 3-D 棒(3)...
 折れ線(L)...
 面(A)...
 円(E)...
 ハイロー(H)...
 箱ひげ図(X)...
 エラーバー(O)...
 人口ピラミッド(Y)...
 散布図/ドット(S)...
 ヒストグラム(I)...

投入済み変数または除去された変数^a

モデル	投入済み変数	除去された変数	方法
1	時間 ^a	.	入力

a. 要求された変数がすべて入力されました。
 b. 従属変数 件数

モデル集計

モデル	R	R ² 乗	調整済み R ² 乗	推定値の標準誤差
1	.959 ^a	.920	.910	7.904

a. 予測値: (定数)、時間。

単純散布図

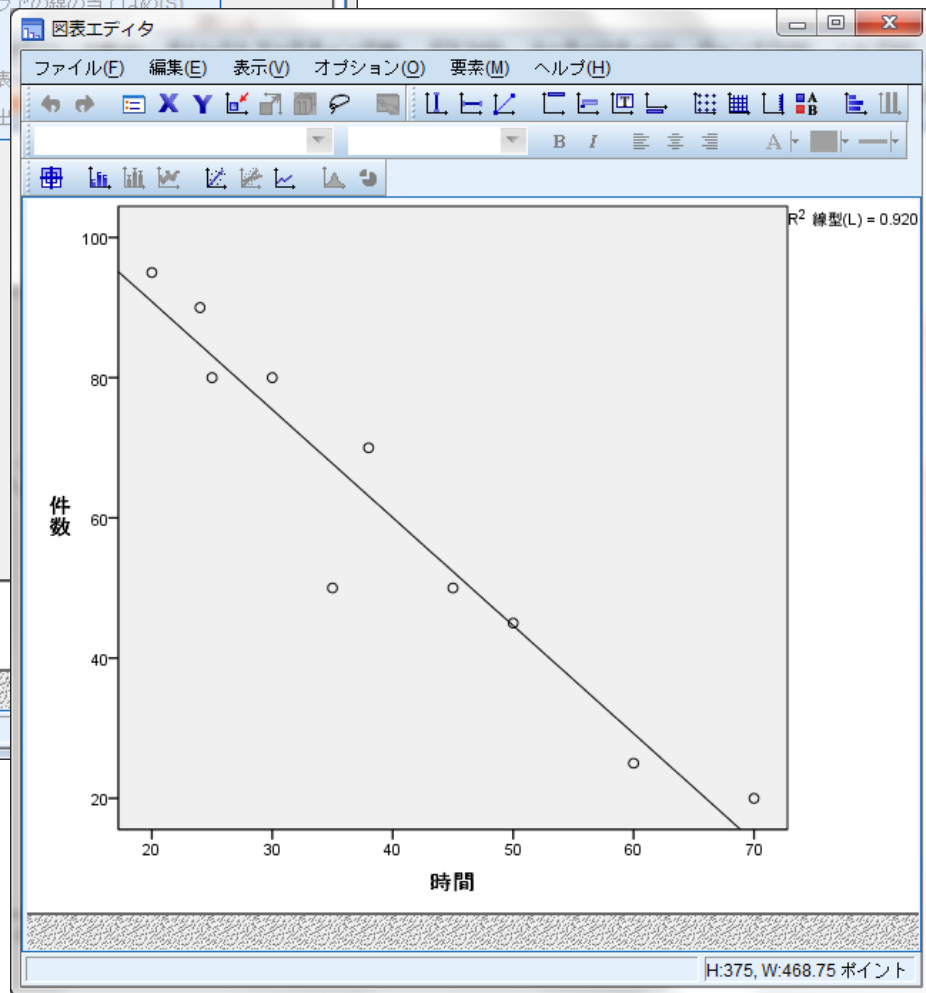
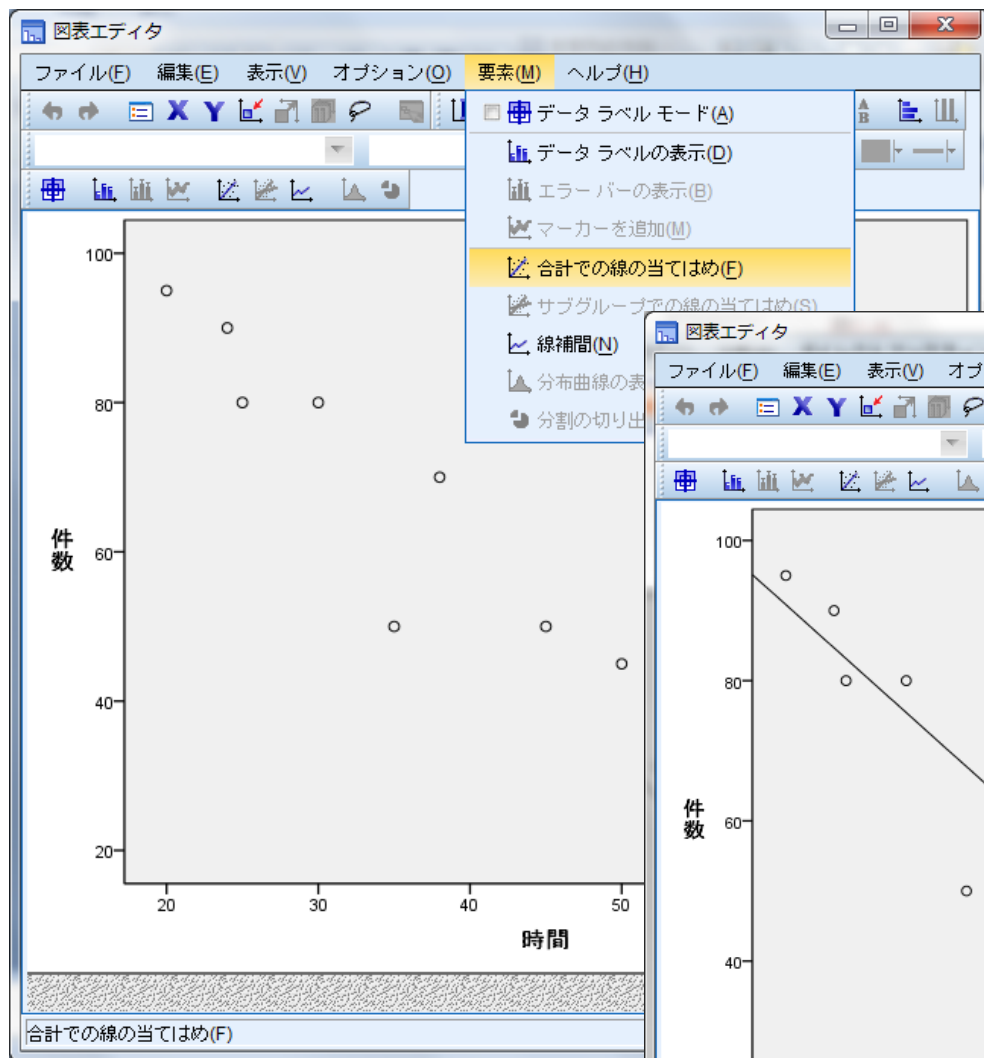
Y 軸(Y): 件数
 X 軸(X): 時間
 マーカーの設定(S):
 ケースのラベル(C):
 パネル
 行(W):
☐ 変数を入れ子にする (空白行なし)(N)
 列(L):
☐ 変数を入れ子にする (空白列なし)(E)
 テンプレート
☐ 指定された図表を使用(U):
 ファイル(F)...
 OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

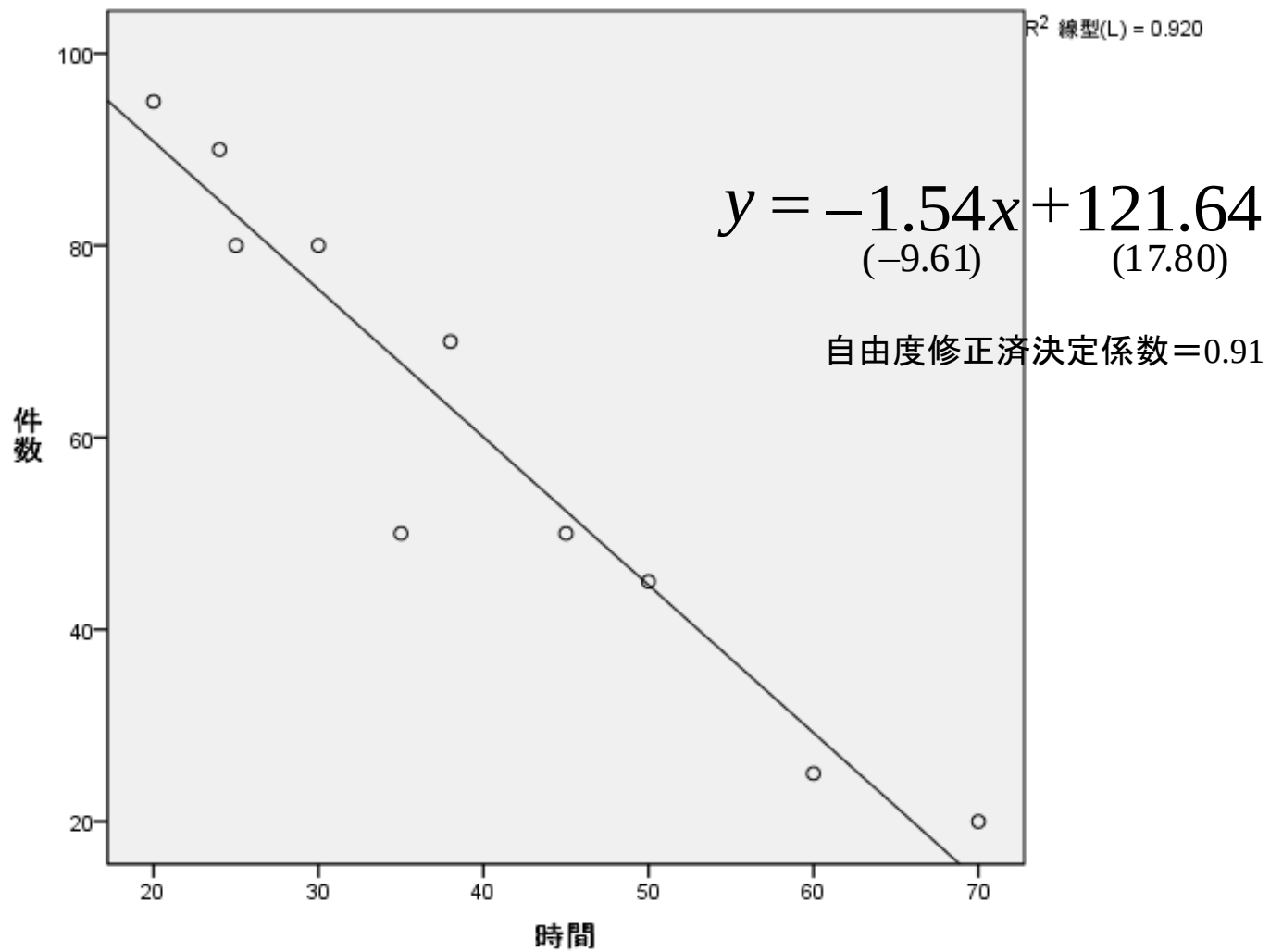
散布図/ドット

単純な散布
 行列散布図
 シンプルドット
 オーバーレイ散布
 3-D 散布図
 定義 キャンセル ヘルプ

IBM SPSS Statistics プロセッサは使用可能です

22:59





【演習】

下記データを用いて単回帰分析を行ってください(回帰線も図式化すること)。

従業員数(独立変数)と売上額(従属変数)の関係

店舗	従業員数 (人)	売上額 (百万円)
A	107	2,861
B	336	8,513
C	233	5,895
D	82	3,897
E	61	1,589
F	378	10,378
G	129	4,636
H	313	5,634
I	142	3,722
J	428	10,201