

データを階級ごとに整理した表を何というか。

【解説】

解説ビデオクリップ



データを階級(Class)ごとに整理した表が、度数分布表(Frequency Distribution Table)である。

例： 31 名のテストの点数

元データ

12	45	63	69	76				45	39
	52		69	86		55	59	59	
65	82	89	73			68	74	64	55
60	40	40	35	60	76	82	79	60	65

度数分布表

階級	人数
10	
20	
30	
40	
50	
60	
70	
80	
90	
100	
計	

元データとして 40 個のマスに 31 個のデータがある。これらを右表のように整理したのが、度数分布表である。対応した階級にいくつのデータがあるのかを示しており、条件にあうようにクラス分けをするというイメージである。作り方で注意すべきことは以下の通りである。

1. 階級数を 10~20 に設定(データ数に応じて設定)
2. 階級の間隔は統一
3. 同じデータが2つの階級に入らないように
4. データ数と各階級の度数の合計が一致することを確認

Excel を利用すれば、100 個以上のデータでも簡単に整理ができる。度数分布表の作成には Frequency 関数を利用するが、配列を利用しなければならない。

【関連問題】

年 月 日

1. 上に示した度数分布表を完成させる。
2. 階級を英語で何というか。
3. Excel 関数で頻度を計算する関数は何であるか。

度数分布表から作成されるグラフを何というか。

【解説】

解説ビデオクリップ



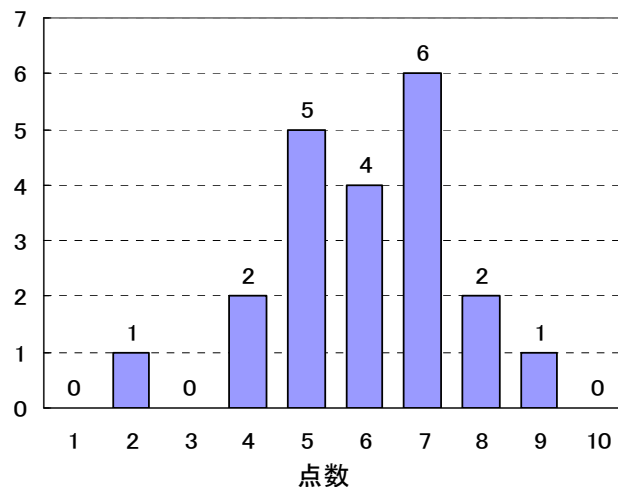
度数分布表をもとにグラフ化したものが、ヒストグラム (histogram) である。縦軸が度数、横軸が各階級幅であり、ヒストグラムの特徴として、各階級の柱の面積が級度数を正確に反映していることが挙げられる。データ数が多く階級幅が細かいときには、折れ線グラフを用いるとデータが連続してよくわかる。

度数分布表

点数	人数
1	0
2	1
3	0
4	2
5	5
6	4
7	6
8	2
9	1
10	0
合計	21

人数

ヒストグラム



ヒストグラムを見るとデータの分布 (distribution) がよく分かる。データの散らばり具合、歪み (skewness)、^{とが}尖り (kurtosis) などが把握できる。

【関連問題】

年 月 日

1. 上のデータで、最高点は何点であるか。
2. 上のデータで、最低点は何点であるか。
3. 上のデータで、最も人数が多いのは何点であるか。

テストの成績が 80 点、55 点、90 点、60 点、65 点であるとき、メジアンは何点か。

【解説】

解説ビデオクリップ

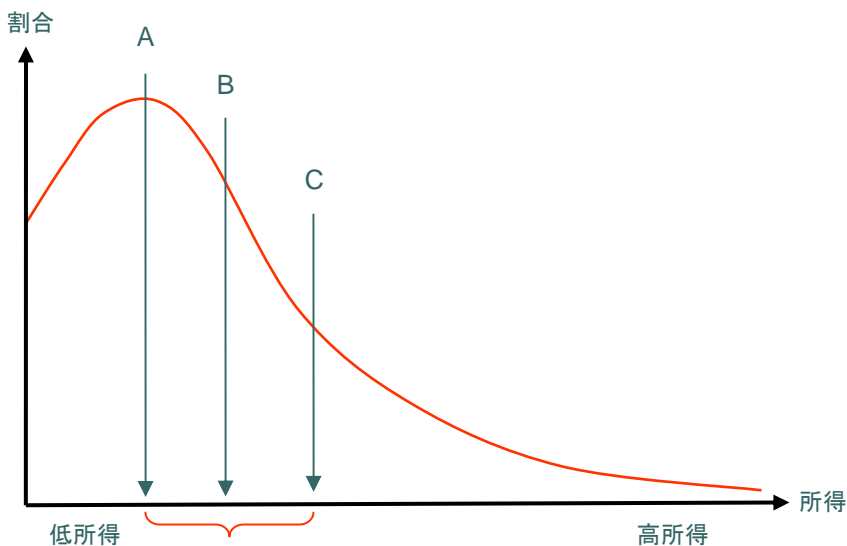


メジアンは中央値であるから、5 つのデータであれば上から 3 番目、下から 3 番目のデータとなる。5 つのデータを並べ替え (Sort) すると、90,80,65,60,55 という順になる (降順)。よってメジアンは 65 点。

メジアン (median) とは、中央値、中位数のことである。Excel では、=median(範囲)を入力する。データの代表値として、他に算術平均 (mean) やモード (mode) がある。最もよく利用されるが平均である。これはデータの総和をデータ数 (n) で割ったものである。平均 (μ) を数式で示すと以下のようになる。

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

日本の家計の貯蓄分布を表現したのが右図である。A はモード、C は平均を示している。この図からわかるように、一部のお金持ちが平均を押し上げているので、平均値は実態よりも大きく感じる。モードを用いる方が日本経済の実態をよく表したデータであろう。



【関連問題】

年 月 日

1. テストの成績が 50 点、65 点、70 点、40 点、75 点であるとき
 - A) 平均は何点か。
 - B) メジアンは何点か。
2. 11 人のデータがある。メジアンは上位何番目のデータを利用すればよいか。

テストの成績が 80 点、55 点、90 点、60 点、65 点であるとき、レンジはいくつか。

【解説】

解説ビデオクリップ



データの散らばりの尺度としてレンジ(範囲, Range)がある。範囲は、**最大値**(maximum)から**最小値**(minimum)を引いたものである。最大値は 90、最小値は 55 なので、レンジは 35(長さ)となる。

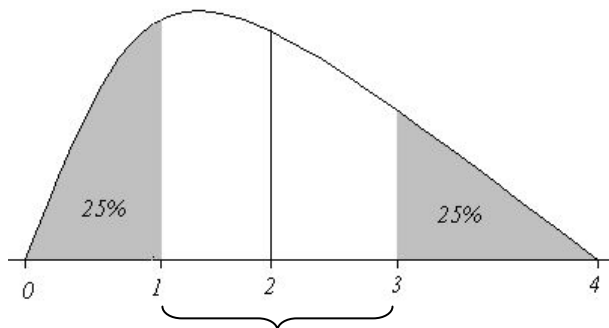
なお、最大値を求める Excel 関数は=MAX、最小値は=MIN である。これはいずれも位置を示している。レンジは、全データ(n 個)のうちで 2 つのデータしか使っていない。すなわち残りのデータ($n-2$ 個)の持つ情報は使われていない。レンジの問題点としては、**外れ値**(^{はず}一つだけとんでもなく異なった値のデータ)の影響を受けやすいことが挙げられる。データ数が大きくなるに従い、レンジも大きくなる傾向があるので注意する。

レンジの問題点を補正する尺度として、**四分位範囲**(interquartile range)がある。これはデータを小さい順番に並べ四分割(各 25%)し、上位 25%の点(**第1四分位**)と 75%の点(**第3四分位**)の範囲を四分位範囲とする。

$$\text{四分位範囲} = \text{第3四分位点} - \text{第1四分位点}$$

Excel の標準関数を使った計算は、=QUARTILE(データ, 3) - QUARTILE(データ, 1)となる。ここで戻り値は 0~4 であり、それぞれ図中の数値を表示する。

- 0: 最小値
- 1: 第1四分位点
- 2: 第2四分位点(中位数:メジアン)
- 3: 第3四分位点
- 4: 最大値



【関連問題】

年 月 日

1. 最大値を求める Excel 関数は何か。
2. 最小値を求める Excel 関数は何か。
3. 上の図で範囲(レンジ)を示しなさい。

テストの成績が 80 点、55 点、90 点、60 点、65 点であるとき、標準偏差はいくつか。

【解説】

解説ビデオクリップ



テストの点数の総和 (summation) は、 $80 + 55 + 90 + 60 + 65 = 350$ である。①平均 (μ) はこれをデータ数 (n) で割ったものだから、 $350/5 = 70$ 点となる。

平均を基準として、どれだけ^{かたよ}偏りがあるかが②偏差 (deviation) である。したがって、 $x - \mu$ を計算する。80-70, 55-70, ..., 65-70 であるから、それぞれ、10、-15、20、-10、-5 となる。プラス・マイナスという符号が混在するので、これを2乗 (③偏差平方) すれば、すべて正の値となる。100, 225, 400, 100, 25 を合計すると 850 となり、これを④偏差平方和と呼ぶ。これをデータ数で割ったものが⑤分散 (variance: σ^2) であるから、170 となる。ただし、これは2次元であるので、これを1次元に戻すために平方根 (square root) を取る。標準偏差 (standard deviation: σ) は分散の⑥平方根なので、約 13.038... となる。

標準偏差の導出には、以上の①～⑥の計算手順が必要である。計算式で表現すると下のようになる。

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$$

この両辺を2乗した σ^2 は分散となる。分散は偏差平方の平均でも求められる。

Excel で標準偏差を求めるには、STDEV 関数を利用する。なお、上記の計算を Excel で実行するには、①=AVERAGE、④=DEVSQ、⑤=VARP、⑥=SQRT を利用する。

自学自習: 統計学入門

【関連問題】

年 月 日

テストの成績が 80 点、75 点、70 点、60 点、65 点であるとき

1. 平均は何点か。
2. 偏差平方和はいくつか。
3. 分散はいくつか。
4. 標準偏差はいくつか。

テストで平均点が 60 点、標準偏差が 15 であった。ある学生は 75 点をとったという、偏差値はいくつになるか。

【解説】

解説ビデオクリップ



偏差値計算は標準化した値(Z)を 10 倍し、これに 50 を加える。すなわち、 $50+10\times(75-60)/15$ となり、偏差値は 60 となる。

標準化の計算には、平均(μ)と標準偏差(σ)を用いる。

$$Z = \frac{x - \mu}{\sigma}$$

標準化された Z は、平均が 0、標準偏差が 1 というスケールになる。Z は小さな値で、符号はプラスとマイナスがあるために、人間の目で数値を比べるには面倒である。そのような不便をなくすためにこれを 10 倍し、50 を加える。偏差値は以下のような計算となる。

$$50 + 10 \times \frac{x - \mu}{\sigma} = 50 + 10 \times Z$$

位置の尺度としての平均と散らばりの尺度としての標準偏差を用いており、平均を 50 とするモノサシである。偏差値では、成績が 1 番でも飛び抜けた 1 番かどうかとも判断できる。最高点がともに 95 点であり、それぞれの偏差値が 70、75 であった場合に、後者の科目の方が他の受験生の平均的能力に比較してより優れていると判断ができる。

Excel では `standerize` 関数を利用すれば、Z が計算できる。

【関連問題】

年 月 日

1. 平均点が 70 点のテストで 70 点をとった学生の偏差値はいくつであるか。
2. 偏差値が 50 であった。これはどのようなことを意味するか。
3. 標準化 Z を 10 倍し、50 を加える操作にはどのような意味があるか。

正規分布のグラフはどのような形状であるか。

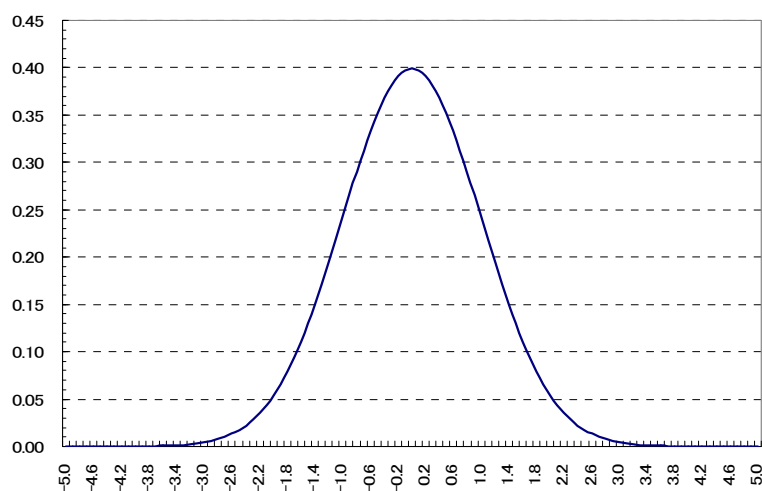
【解説】

解説ビデオクリップ



正規分布は Normal Distribution といい、以下の図のような平均 μ を中心として左右対称 (symmetry) の釣り鐘のような形をしている。平均値に近い値は高く、平均から離れるほど低くなる。散らばりの尺度である標準偏差 σ によって、その形状は変化する。すなわち、 σ が大きければ山は低くなり、裾野が広がる。また、この曲線はガウス (Gauss) 曲線とも呼ばれる。

標準正規分布とは平均が 0 で標準偏差が 1 という場合であり、 $N(0,1)$ と表現する。以下は標準正規分布である。



正規分布の密度関数は以下の式の通りである。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

標準正規分布を Excel で表現すると =EXP(-(X^2)/2)/SQRT(2*PI()) になる。

【関連問題】

年 月 日

1. 平均 0、標準偏差 1 である正規分布を何というか。
2. $N(0,1)$ とはどのような意味か。
3. Excel で円周率を計算するには、どのような関数を用いるか。

データ x とデータ y の相関関係を調べる場合、どのようなグラフを作成すればよいか。

【解説】

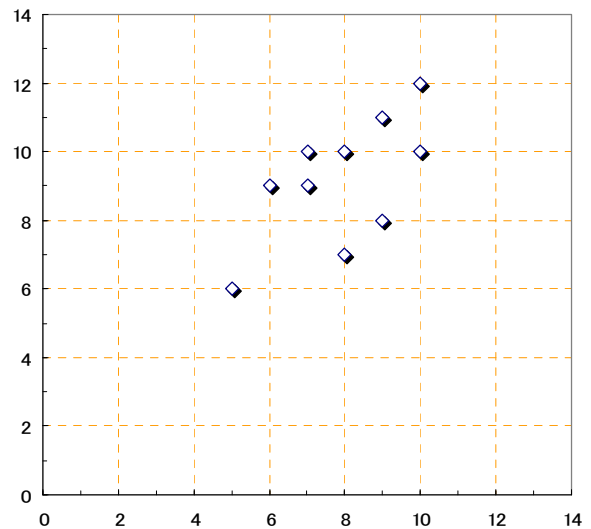
解説ビデオクリップ



2 種類のデータの関係を調べるには、横軸に x 、縦軸に y をとった 散布図 (scatter diagram) が適当である。Scatter とは「ばら撒く」という意味である。散布図をプロット図ともいう。プロット (plot) とは ^{びょう} 鋌を打つようなイメージである。すなわち、 xy 平面にデータに対応する点を打つことである。

標本番号	x	y
1	9	11
2	7	10
3	10	12
4	5	6
5	8	10
6	8	7
7	6	9
8	7	9
9	9	8
10	10	10
合計		
平均		
標準偏差		

散布図



データ x の平均を \bar{X} で表し、データ y の平均を \bar{Y} で表す。

【関連問題】

年 月 日

1. plot とはどのような意味か。
2. scatter とはどのような意味か。
3. 上の表の空欄(合計、平均、標準偏差)を計算する。
4. 上の散布図に x と y の平均値を書き込む。

相関係数の分母に利用されているのはどのような式か。

【解説】

解説ビデオクリップ(TIES)



以下の式はピアソンの積率相関係数 (correlation coefficient) である。

$$r = \frac{\sum (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum (x_i - \bar{X})^2 \sum (y_i - \bar{Y})^2}}$$

分母・分子ともに xy の各偏差が利用されている。分母は x の偏差平方和と y のそれで構成され、さらにそれらの積に対し平方根をとったものである。

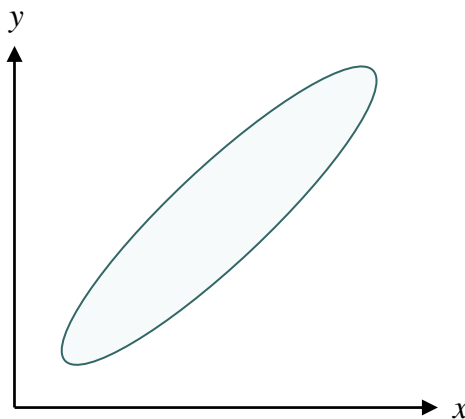
分母は平方根がついており、必ず正の値となる。よって、分子である xy の偏差の積和がこの係数 r の符号を決定する。また、この式から相関係数は1から-1の値をとることがわかる。

$$-1 \leq r \leq 1$$

r の値で以下のように判断する。

- $r > 0$: 正の相関
- $r < 0$: 負の相関
- $r = 0$: 無相関
- $r = \pm 1$: 完全相関

右図は、 x と y が正比例している正の相関を表している。
また、完全相関の場合には、データが一直線上に並ぶ。



Excel で相関係数を計算するには CORREL 関数を利用する。

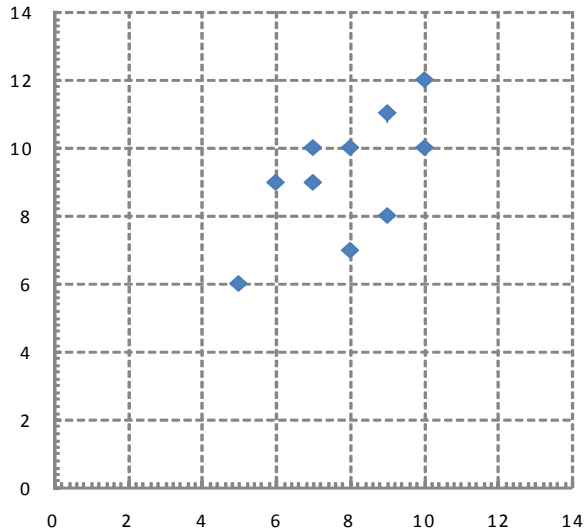
【関連問題】

年 月 日

1. 上の相関係数 r の計算式で、 y の偏差平方和を表している箇所を丸で囲む。
2. 相関係数が 1 であった。そのときの散布図はどのようなものであるか。
3. 前の設問にある x と y のデータの相関係数を計算する。

以下の 2 つのデータを元に作成した散布図がある。図中に☆マークで平均を、中心を通るような直線をグラフに記入しなさい。また、直線の y 切片の値を求めなさい。

標本番号	x	y
1	9	11
2	7	10
3	10	12
4	5	6
5	8	10
6	8	7
7	6	9
8	7	9
9	9	8
10	10	10
合計		
平均		



【解説】

解説ビデオクリップ



表にある 10 個のデータを xy グラフにプロットしたのが上のグラフである。横軸が x データで縦軸が y データである。これらの平均 (7.9, 9.2) を散布図に☆のマークで記入をする。

平均はデータの中心を示す一つの指標であるから、中心を通る直線は平均☆を通るように記入する。直線は以下のような式で表現される。

$$Y = a + bX$$

ここで a は直線の Y 切片であり、 b は傾きである。

直線の切片 a は、縦軸 Y との交点で測る。グラフに補助目盛りをつけると求めやすい。直線の傾き b は、 Y の増加量 / x の増加量で測定する。たとえば、 x の増加量を 10 としたときに、 Y がどれだけ増加したかを計算すればよい。

$$b = \frac{\Delta y}{\Delta x} = \frac{y \text{ の増加量}}{x \text{ の増加量}}$$

【関連問題】

年 月 日

1. 各自が記入した直線の傾き b はおおよそいくつになるか調査せよ。
2. $Y = a + bX$ の a, b に各自のデータを当てはめ、 $X=10$ であるときの Y の値を求めよ。

中心を通ると思われる直線とデータの差を何というか。

【解説】

解説ビデオクリップ



中心を通ると思われる直線とデータの差を残差 (residual)あるいは誤差 (error、略語は e)と呼ぶことがある。データを現実の値とすれば、直線上の点は推定された値となる。ここでは縦方向(データ Y)のズレに注目する。その関係を式で表現すれば、以下のようなになる。

$$\text{現実値} - \text{推定値} = \text{推定誤差 (残差)} e$$

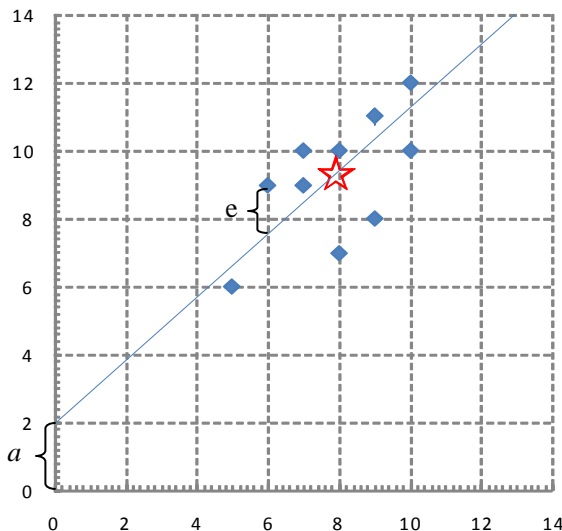
推定誤差 e を 0 に近づけることは、推定誤差が小さくなることなので、中心を通る直線の基準として望ましい。そこで、各推定誤差の和 ($\sum e$) は、以下のような S で表すことができる。

$$S = e_1 + e_2 + e_3 + e_4 + e_5 + e_6 + e_7 + e_8 + e_9 + e_{10}$$

ただし、この式の右辺にある各 e の符号は正と負まちまちであるので、それを統一するために各 e を 2 乗し、符号をプラスに揃える。推定誤差を平方した値 (e^2) の合計 ($\sum e^2$) は次のようになる。

$$S^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2 + e_7^2 + e_8^2 + e_9^2 + e_{10}^2$$

この S^2 を 残差平方和 (residual sum of squares) という。




【関連問題】

年 月 日

1. 先の設問で $a=2, b=1$ のとき、10 個のデータの残差を求めなさい。
2. そのときの残差の合計 ($\sum e$) を求めなさい。
3. そのときの残差平方和 ($\sum e^2$) を求めなさい。

残差平方和を最小にする a, b の組み合わせの公式を求めよ。

【解説】

解説ビデオクリップ 

誰が見てもデータ群の中心を通ると思われる直線を引くには客観的な基準が必要である。残差平方和 S^2 を最小にすることは推定の誤差が最小になることなので、中心を通る直線を求めるひとつの基準として望ましい。残差平方和が最小になるように a, b を決める方法を**最小二乗法 (Ordinary Least Square Method: OLS)**という。残差平方和が最小となる公式は以下の通りである。

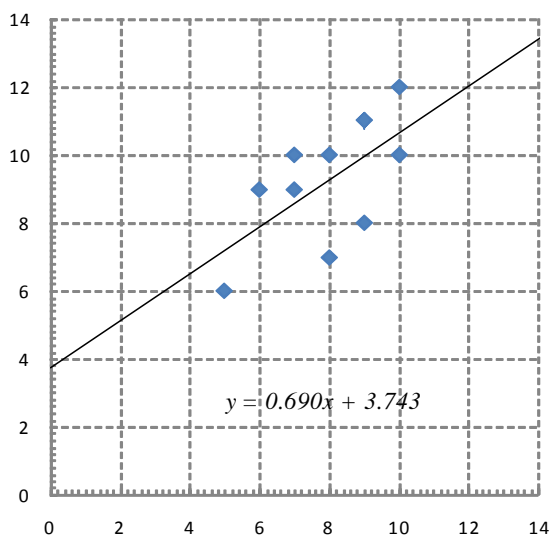
$$a = \bar{Y} - b\bar{X}$$

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

ここで \bar{X} と \bar{Y} はそれぞれデータ x と y の平均値を表している。 a, b を求める計算は以下の手順で行う。まずデータ x, y から b を求め、この b と平均 \bar{X} と平均 \bar{Y} から a を求めればよい。この計算にはデータから平均の差である**平均偏差**が利用されていることに注意する。

残差平方和が最小となる a, b の組み合わせはひとつしかなく、そのとき残差の合計は 0 になる。最小 2 乗推定量の導出には数学的な操作が必要であるが、これについては計量経済学の標準的なテキストを参照されたい。

なお、Excel では右図のように散布図へ近似曲線を挿入できる。



【関連問題】

年 月 日

1. 残差平方和が最小となる a, b はいくつになるか。
2. そのときの残差平方和はいくつになるか。
3. そのときの残差の合計はいくつになるか。

打率 3 割の打者が、1 試合で 3 回の打席が回ってくるものとする。

1. 3 打席連続でヒットを打つ確率
2. 3 打席すべて凡退
3. 少なくとも 1 回、ヒットを打つ確率(正答率:60%)

【解説】

解説ビデオクリップ



1. 3 打席連続でヒットを打つ確率

$$0.3 \times 0.3 \times 0.3 = 0.027 = \underline{2.7\%}$$

2. 3 打席すべて凡退の確率

$$0.7 \times 0.7 \times 0.7 = 0.343 = \underline{34.3\%}$$

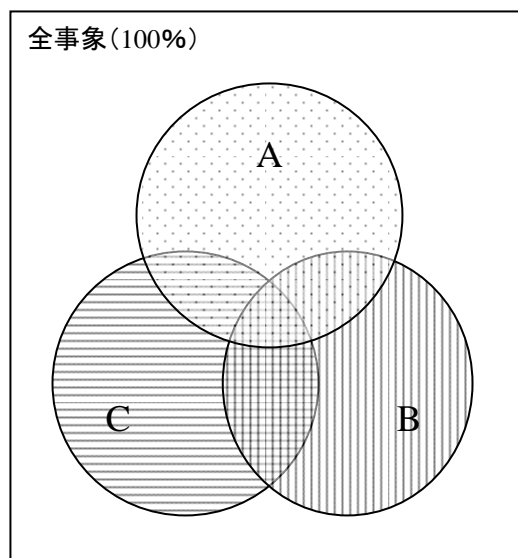
3. 少なくとも 1 回、ヒットを打つ確率

$$1 - 0.343 = 0.657 = \underline{65.7\%}$$

確率(probability)とはその事象(event)が生じる可能性を「たくさんある」、「ほとんどない」というような表現でなく、数値で表す。数値による基準で確からしさを測定する。

全事象を 1(100%)とすると、事象が生じる割合(0%~100%)が確率である。高校までに学習した集合を使えば、図のようになる。全体集合と部分集合の面積の割合として表現する。

例題にある 3 打席連続ヒットは 3 つの円(A,B,C)が重なった部分($A \cap B \cap C$)の面積を表している。 \cap (キャップ)は積事象といい、重なった部分を示す。一回でもヒットを打つという事象は 3 つの円にある($A \cup B \cup C$)部分である。 \cup (カップ)は和事象という。これを計算するには、全打席凡退した事象ではない事象の計算をする。これを排反事象という。上の図(ベンズ)では、全事象(1)から網掛けの面積(0.657)を引けばよい。



【関連問題】

年 月 日

打率 3 割の打者が、1 試合で 4 回の打席が回ってくるものとする。

1. 4 打席連続でヒットを打つ確率
2. 4 打席すべて凡退
3. 少なくとも 1 回、ヒットを打つ確率

Excel の3Dグラフの等高線を利用して、以下の効用関数の3次元グラフ(効用曲面)を描く。

$$U = x_1^{0.3} x_2^{0.3}$$

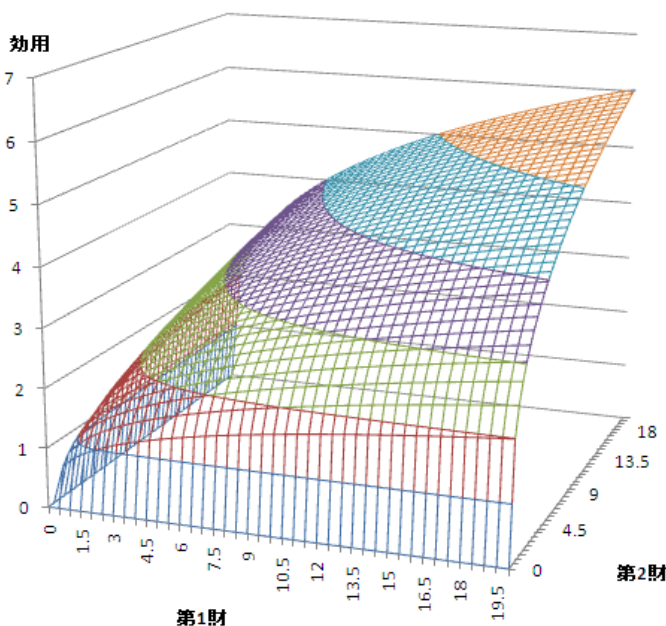
2財(x_1, x_2)から得られる効用データとグラフを完成させる。完成したグラフをさまざまな角度から眺める。

【解説】

解説ビデオクリップ 

Excel の第 A 列に 0 から 0.5 きざみで 20 までの値(第1財: x_1 の数)をとる(縦 2~41 行まで)。次に、第 1 行にも同じ幅で値(第2財: x_2 の数)をとる(横 B~AP 列まで)。関数の式をひとつのセル(例えば B2)に書き込み(B2 への入力 :=(\$A2^0.3)*(B\$1^0.3)と書く)計算する。それを縦横に連続コピーして、B2 から AP41 まで 1600(40×40)個作成する。計算の際に\$マークをつけているのは、連続コピー(オートフィル)をしても参照元が移動してしまわないように、列や行を固定するためである。

1600 個のデータを元にして等高線グラフを利用すれば、滑らかな3次元(3 dimension)の効用曲面を描くことができる。等高線グラフは地図の等高線と同じ概念である。等高線グラフの元になるデータ高さを表しており、この例題では効用水準(U)である。この図を横から眺めると効用曲線(utility curve)が、上から眺めると無差別曲線(indifference curve)が現れる。



【関連問題】

年 月 日

1. 上の例題で作成した効用曲面において指数の値を 0.8 と 0.2 とする場合の曲面を描く。
2. 上の例題と比べて、グラフのどこが異なるか。
3. この関数を何というか。(コブ・ダグラス型関数)

以下の表は、Excel で作成した牛の生産業者の生産と費用の関係表(単位は、牛の生産量:頭数、費用:万円)である。この表の空欄を完成させ、固定費用を示す。

	A	B	C	D	E
1	生産量	総費用	可変費用	平均費用	限界費用
2	0	250	0	0	150
3	1	400	150	400.00	110
4	2	510	260	255.00	90
5	3	600	350	200.00	80
6	4	680	430	170.00	
7	5	754	504		
8	6	824	574		
9	7	892	642		
10	8	960	710		
11	9	1030	780		
12	10	1103	853		
13	11	1180	930		
14	12	1262	1012		
15	13	1350	1100		
16	14	1445	1195		
17	15	1548	1298		
18	16	1660	1410		
19	17	1784	1534		
20	18	1924	1674		
21	19	2086	1836		
22	20	2278	2028		-


【解説】

解説ビデオクリップ 

生産者は、財を生産・販売して得られる利潤が最大となるように行動する。利潤(Profit)は、収入(Revenue)－費用(Cost)で表せる。ここでは、費用と生産量の関係を表す費用関数(cost function)が数値で示されている。B列にある総費用やC列の可変費用を折れ線グラフで描くと3次曲線として表現される。これらの曲線は、 $C = ax^3 + bx^2 + cx + d$ のような3次関数となる。総費用は可変費用に250だけ増やしたものであることがわかる。ここでの 250 は固定費用という。

表で空欄となっている平均費用(average cost)は総費用を生産量で割ったものであるなので、セルD7の平均費用の計算は、 $=B7/A7$ となる。限界費用(marginal cost)は生産量を1増やした時に増加する費用なので、セルE6の限界費用の計算は、 $=C7-C6$ となる。それぞれをオートフィルで連続コピーする。

この2種類のデータ(D3～E21)をグラフ化すれば、2次関数の曲線となる。

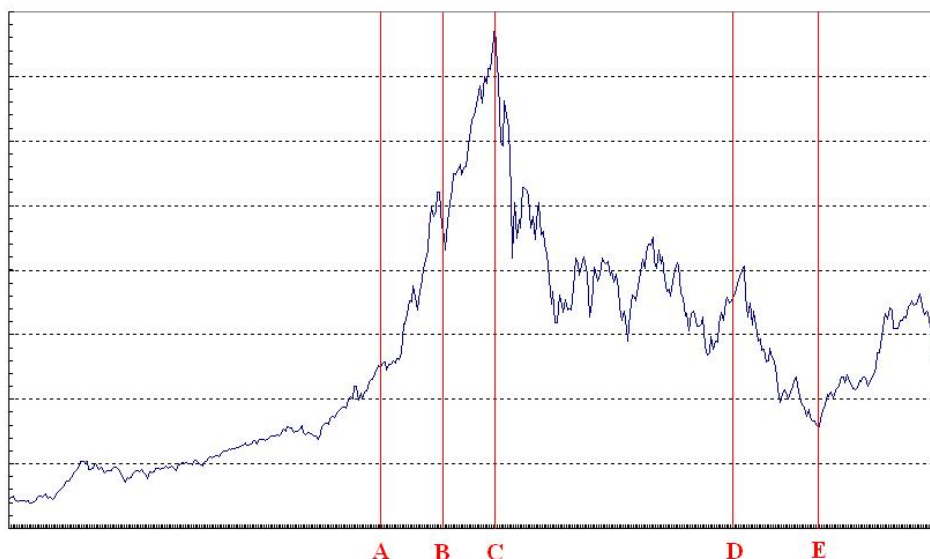
 クロス参照:#048

【関連問題】

年 月 日

1. 折れ線グラフで総費用を描く。
2. 折れ線グラフで可変費用を描く。総費用とどこが異なるか。
3. 折れ線グラフで平均費用と限界費用(D3～E21を範囲指定)を同時に描く。

以下のグラフは日経平均株価(1972年-2008年)の推移を表している。「ブラックマンデー」と呼ばれる時期はA～Eのどれか。



【解説】

解説ビデオクリップ



日本経済に関するデータは、日本銀行のサイト(<http://www.boj.or.jp/>)から入手できる。インターネット上の経済データを元に Excel の折れ線グラフで表現すると経済の動きが明らかになる。例題のグラフは、日経平均株価(月次データ)を利用し、横軸の左端は 1970 年で、右端は 2008 年ある。ブラックマンデーは日本の株価が最高値になる 2 年前(1987 年 10 月 19 日 月曜日)に起こり、グラフでは B の位置 である。史上最大規模の世界的株価の暴落であった。日経平均株価が最も高かったのはグラフでは C の位置であり、1989 年 12 月 29 日の 38,915 円である。グラフの横軸 A は 1985 年頃であり、プラザ合意により円高が始まった頃でもある。E の位置は小泉純一郎内閣が「構造改革」に着手し、不良債権処理などによって株価にも影響を与えたことによる。

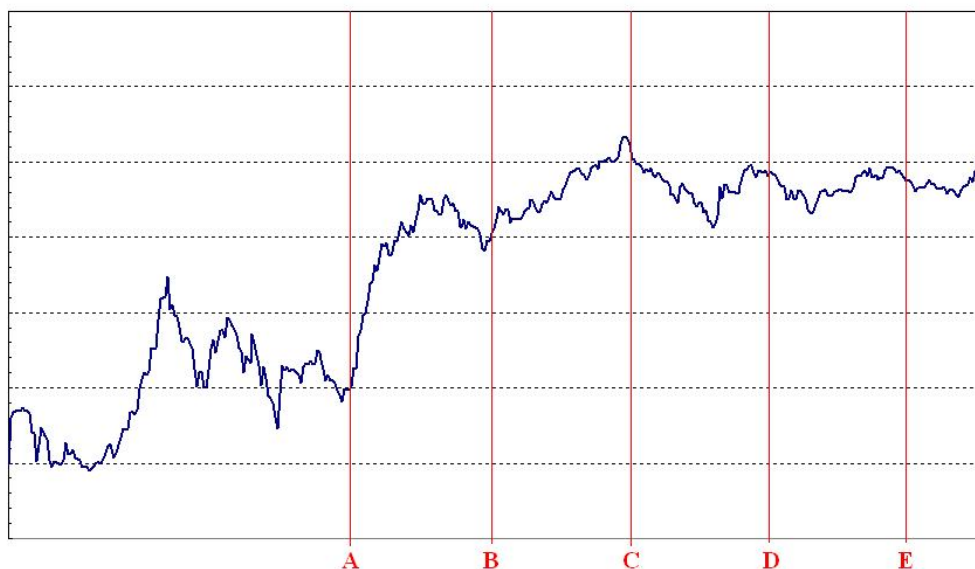
👉 クロス参照: #133,297

【関連問題】

年 月 日

1. 日経平均株価のデータは、どのサイトから入手できるか。
2. 日経平均株価が最も高かったのはいくらか。
3. 上のグラフにある縦軸にメモリアル(値と単位)を記入する。

以下のグラフは円のドルに対する為替レートの推移(1973年-2008年)を表している。プラザ合意はA～Eのどの時期であるか。



正答率: 70%

【解説】

解説ビデオクリップ



この経済データのグラフは、円とドルの交換レート(exchange rate)で1ドル何円であるかを表している。縦軸の一番上は0円、一番下は350円である。通常のグラフとは逆(Excelで軸の反転)になっているのは、グラフの上方への動きが円高ということを示すためである。グラフは月次データを利用しており、横軸の左端は1973年で右端は2008年ある。1971年のニクソンショックまで固定相場制の下、1ドル360円であったが、変動為替相場制へ移行した。プラザ合意は1985年9月22日(グラフのAの位置)にニューヨークのプラザホテルで開催されたドル安を容認した会議を指す。最も円高であったのは、1995年4月19日に記録した79円75銭が最高値である。

👉 クロス参照:#146,147


【関連問題】

年 月 日

1. 上のグラフの縦軸にメモリアル(値と単位)を記入する。
2. 2008年までに1ドルに対して円の最高値はいくらであったか。
3. 急激な円高は日本経済にどのような影響を与えたか。(＃147参照)

Excelでは表のようなシートで計算処理ができる。Excel2003の1枚のシートには、全部でいくつのセルがあるか。それは2の何乗か。

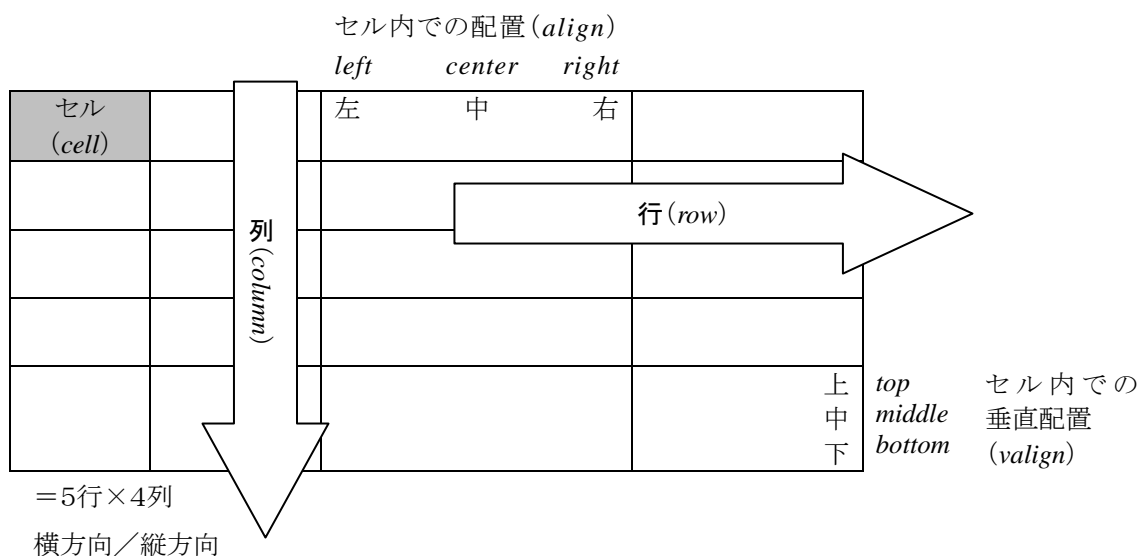
【解説】

解説ビデオクリップ 

Excelを表計算ソフトといい、**スプレッドシート** (spread sheet) 内でさまざまな計算ができる。Excel2003の1枚のシートの中に行 (row) は $65,536 (=2^{16})$ 、列 (column) は $256 (=2^8)$ があるので、**セル (cell)** は $65,536 \times 256 = 16,777,216 (=2^{16} \times 2^8 = 2^{16+8} = 2^{24})$ となる。Excelのシートには行番号 (数字) と列番号 (英字) がついており、この番号を参照してセル番地が一意に決定される。

以下の表は5行4列で構成されており、マスであるセルは20個ある。

■行と列



Excelではさまざまな計算ができ、数学 (円周率、指数、対数など) や統計 (平均、標準偏差、分布など) の関数が用意されている。その他、文字列操作や日付時刻なども計算することができ、活用できる場面は多い。例えば、日付計算は金利計算などに応用できる。

【関連問題】

年 月 日

1. Excel で 2^{16} を計算する。
2. Excel で $\text{pi}()$ を計算する。
3. Excel を利用して、今日までの生存日数を計算する。

457 兆円という日本の GDP データを桁区切り、通貨スタイル、指数表示で示す。

【解説】

解説ビデオクリップ



457 兆を桁区切り (comma separated) 表現にすれば、457,000,000,000,000 という表記になる。このように桁数が大きくなると分かりづらくなるので 3 桁の区切りで表現するとよい。また、通貨スタイルにすると ¥457,000,000,000,000 のようになる。

桁数や小数点以下のそれが大きかったりする場合、Excel では指数表現を用いることがある。例えば、10000 を 1.00E+05 というように 0 を省くような表現ができ、457 兆は 4.57E+14 という表現になる。E の後にある符合が + は桁数が大きいケース、- は小数点を表現する。

スタイル名	入力データ	セル表現
桁区切り	4570000000000000	457,000,000,000,000
通貨スタイル	4570000000000000	¥457,000,000,000,000
指数	4570000000000000	4.57E+14
パーセント小数	0.03	3%
小数点 (3桁)	3.14159265	3.14
小数点 (7桁)	3.14159265	3.1415927

Excel を扱う際の注意として、セルに「10E5」と入力すると「10000」と認識して「1.00E+05」と表示される。また、「10/1」や「10-1」と入力すると日付と判断し、「10 月 1 日」と表示される。

Excel では 1900 年 1 月 0 日 (このような日付は存在しない) を基準 0 として、日付の計算を行う。日付と時刻が基準値からカウントした値をシリアル値 (serial value) としている。日付を整数部、時刻を小数部として表現する。時刻は 24 時間を 0~1 の値で割り当てるので、例えば、0.25 は午前 6 時になる。

40087.5 というシリアル値を日時表現にすれば「2009/10/1 12:00:00」となる。

【関連問題】

年 月 日

- 3.5E-08 はどのような数値の指数表現であるか。
- 1 を日付表示にすると、何年何月何日になるか。
- 37145 を日付表示にすると、何年何月何日になるか。
- 0.5 を時刻表示にすると、何時何分になるか。