

ギリシア文字で μ 、 δ 、 σ を何と読むか。

【解説】

□解説ビデオクリップ



ギリシア文字は、経済学や数学・統計学などの計算における記号として利用することがある。大文字・小文字があるが、これらの読み方は以下の通り。 μ 、 δ 、 σ はすべて小文字である。

表 ギリシア文字の読み一覧

大文字	小文字	読み方	記号として利用
A	α	アルファ	係数
B	β	ベータ	係数
Γ	γ	ガンマ	係数
Δ	δ	デルタ	微分
E	ε	イプシロン	
Z	ζ	ゼータ	
H	η	エータ	弾力性(#036)
Θ	θ	シータ	角度
I	ι	イオタ	
K	κ	カッパ	
Λ	λ	ラムダ	
M	μ	ミュー	平均
N	ν	ニュー	
Ξ	ξ	クシー	
O	$\omicron$	オミクロン	
Π	π	パイ	総積、円周率、利潤(#039)
P	ρ	ロー	密度
Σ	σ	シグマ	総和、標準偏差
T	τ	タウ	(#147)
Υ	υ	ウプシロン	
Φ	ϕ	ファイ	直径
X	χ	カイ	χ^2 分布
Ψ	ψ	プサイ	
Ω	ω	オメガ	オームの単位記号

【関連問題】

年 月 日

1. マクロ経済学・ミクロ経済学で利用されたギリシア文字をチェックする。
2. 円周率を表すには、どのようなギリシア文字が利用されるか。
3. 平均を表すには、どのようなギリシア文字が利用されるか。
4. 標準偏差を表すには、どのようなギリシア文字が利用されるか。

関数 $y = f(x)$ を微分するとは、どのようなことか。

【解説】

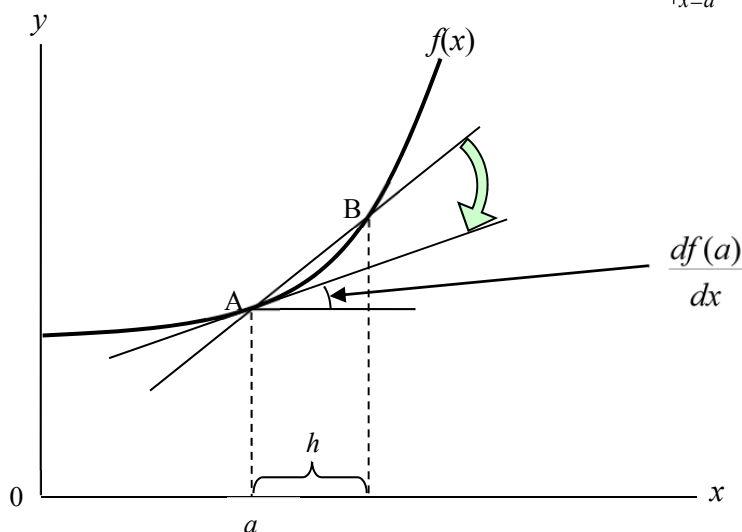
□解説ビデオクリップ



$x=a$ において $y=f(x)$ のグラフの傾き(=接線の傾き)を求める。図のようにグラフ上に点 A、B をとると、直線 AB の傾きは $\frac{f(a+h)-f(a)}{h}$ である。ここで、B を A に限りなく近づけると、グラフが滑らかならば、

接線の傾きは $\lim_{h \rightarrow 0} \frac{f(a+h)-f(a)}{h}$ で与えられる。極限值が存在する時、 $f(x)$ は $x=a$ において微分可能

であるといい、この極限値を a における $f(x)$ の微分係数と呼び、 $\frac{df(a)}{dx}$, $\left. \frac{df(x)}{dx} \right|_{x=a}$, $f'(a)$ などと表す。



$f'(a)$ の a は任意の点だから、 x と書き改めると、 $f'(x)$ はグラフ上の点 $(x, f(x))$ における曲線の傾きを与える関数であり、導関数と呼ばれる。これを $\frac{dy}{dx}$, $\frac{df(x)}{dx}$, y' , $f'(x)$ のように表す。

関数の導関数を求めることを、「関数を微分する」(differentiate)という。次の公式は、特に重要である。

$$(x^n)' = nx^{n-1}$$

$$\{cf(x)\}' = cf(x)' \quad [c \text{ は定数}]$$

$$\{f(x) \pm g(x)\}' = f'(x) \pm g'(x) \quad \{f(x) \cdot g(x)\}' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

定義域の内部において $f(x)$ が極大値または極小値(まとめて極値という)をとる x では、 $f'(x) = 0$ となることを利用して、関数の最大値または最小値を求めることができる。

2変数の関数 $z = f(x, y)$ のケースで偏微分の意味を説明し、極値の必要条件について述べよ。

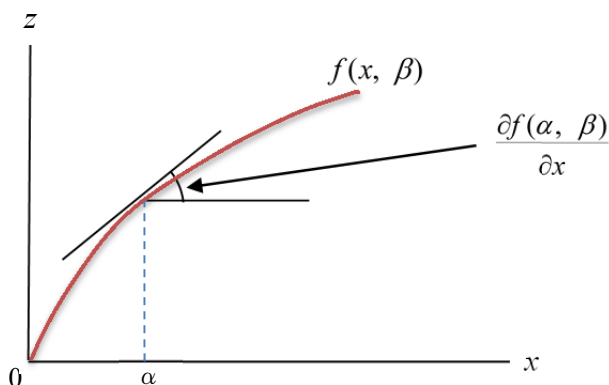
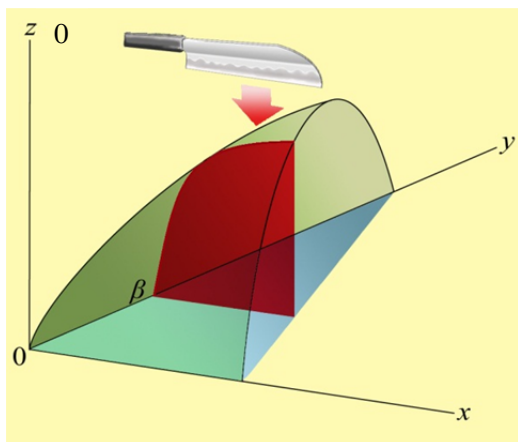
【解説】

□解説ビデオクリップ



$z = f(x, y)$ のグラフを (α, β) を通り、 x 軸に平行に真上から切ると、切り口は $z = f(x, \beta)$ のグラフとなる。そのグラフの $x = \alpha$ における接線の傾きを x に関する偏微分係数と呼び、 $\frac{\partial f(\alpha, \beta)}{\partial x}$ のように表す。

同様に、 y に関する偏微分係数を $\frac{\partial f(\alpha, \beta)}{\partial y}$ のように表す。



$\frac{\partial f(\alpha, \beta)}{\partial x}$ で (α, β) は任意の点だから、 (x, y) と書き改め、これを x に関する偏導関数という。 $\frac{\partial f(x, y)}{\partial x}$ を求めることを「 $f(x, y)$ を x で偏微分する」という。

偏導関数 $\frac{\partial f(x, y)}{\partial x}$ を求めるには、 y を所与として x だけの関数とみなして微分すればよい。 y に関する

偏導関数 $\frac{\partial f(x, y)}{\partial y}$ も同様である。(例) $z = x^3 y^2$ のとき、 $\frac{\partial z}{\partial x} = 3x^2 y^2$ 、 $\frac{\partial z}{\partial y} = 2x^3 y$

内点解 (定義域の内部での解) を仮定すれば、 (x^*, y^*) で極値をとるためには、 x 軸・ y 軸のいずれの方向に見ても極大または極小になっているから、

$$\begin{cases} \frac{\partial f(x^*, y^*)}{\partial x} = 0 \\ \frac{\partial f(x^*, y^*)}{\partial y} = 0 \end{cases}$$

を満たすことが必要である。

2変数の関数 $z = f(x, y)$ が制約条件 $g(x, y) = 0$ のもとで極値をとるための必要条件について述べよ。

【解説】

□解説ビデオクリップ



極値の必要条件は、以下のラグランジュ乗数法により機械的に導出できる。

ステップ 1) ラグランジュ関数 L をつくる。

$$L = f(x, y) + \lambda g(x, y)$$

※ λ はラグランジュ乗数と呼ばれる。

ステップ 2) ラグランジュ関数 L を x 、 y 、 λ に関して偏微分してゼロと置く。

$$\begin{cases} \frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 & \cdots \cdots \cdots \text{①} \end{cases}$$

$$\begin{cases} \frac{\partial L}{\partial y} = \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 & \cdots \cdots \cdots \text{②} \end{cases}$$

$$\begin{cases} \frac{\partial L}{\partial \lambda} = g(x, y) = 0 & \cdots \cdots \cdots \text{③} \longrightarrow \text{制約条件が導出される} \end{cases}$$

ステップ 3) ①～③を x 、 y 、 λ について解く。 $(\lambda$ の値を問題にすることは少ない)

実際に必要条件①～③を解くには、①、②より

$$\frac{\partial f}{\partial x} = -\lambda \frac{\partial g}{\partial x}$$

$$\frac{\partial f}{\partial y} = -\lambda \frac{\partial g}{\partial y}$$

と移項し、これらを辺々割り算すれば、 λ が消去されて x, y だけの式

$$\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = \frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}} \quad \cdots \cdots \cdots \text{④}$$

を得る。④式と制約式③を連立させて解を求めればよい。

ラグランジュ乗数法は、中級レベル以上の経済理論で頻繁に利用される数学テクニックであるから、習得しておく必要がある。ただし、ラグランジュ乗数法の正確な理解は、「経済数学」で補うこと。

等比数列 $a, ar, ar^2, \dots, ar^{n-1}$ の和、すなわち

$$S_n = a + ar + ar^2 + \dots + ar^{n-1} \text{ ----- ①}$$

を求めよ。また、 n が無限大となる無限等比数列の和はどのようなになるか。

【解説】

□解説ビデオクリップ



上の等比数列で a を初項、 r を公比と呼ぶ。 n 項の和である①式と、①式の両辺に r を乗じた式を次のように上下に並べて書く。ただし、②式の右辺は微妙にずらして書かれている。

$$S_n = a + ar + ar^2 + \dots + ar^{n-1} \text{ ----- ①}$$

$$rS_n = ar + ar^2 + \dots + ar^{n-1} + ar^n \text{ ----- ②}$$

①式から②式を辺々引き算して、

$$(1-r)S_n = a - ar^n = a(1-r^n) \text{ ----- ③}$$

③式より

$$S_n = a + ar + ar^2 + \dots + ar^{n-1} = a \frac{(1-r^n)}{1-r} \text{ ----- ④}$$

①式で足し算される項の数が無限、すなわち $n \rightarrow \infty$ の場合の和 S は、 $|r| < 1$ ならば、④式で r^n はゼロになるから、

$$S = a + ar + ar^2 + \dots + ar^{n-1} + \dots = \frac{a}{1-r} \text{ ----- ⑤}$$

⑤式は、 a を投資の増分、 r を限界消費性向と解釈すれば、マクロ経済学における乗数理論を意味している。

👉 #007 乗数の計算

👉 #062 信用創造

【関連問題】

年 月 日

- 毎年 1 万円の利子を永久に払い続ける債券(コンソル債券)がある。割引率が 2% のとき、この債券の市場価格を計算しなさい。

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
0～10	
11～20	
21～30	
31～40	
41～50	
51～60	
61～70	
71～80	
81～90	
91～100	
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 関数を利用すればよいが、これは配列数式である。配列数式の入力時には、**Ctrl** キー + **Shift** キー + **Enter** キーで確定する。

【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。
4. Excel で配列数式を入力する時の注意は何であるか。

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



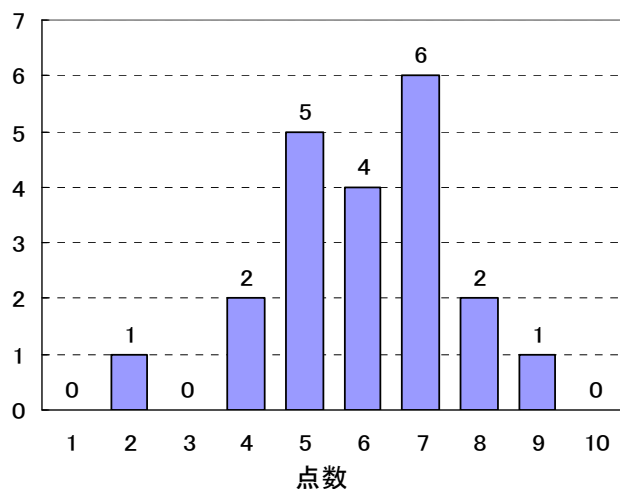
度数分布表をもとにグラフ化したものが、ヒストグラム(histogram)である。縦軸が度数、横軸が各階級幅になっている。ヒストグラムの特徴として、各階級の柱の面積が級度数を正確に反映していることが挙げられる。データ数が多く階級幅が細かいときには、折れ線グラフを用いるとデータが連続してよくわかる。

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、^{ゆが}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、階級はいくつであるか。
2. 上のデータで、最高点は何点であるか。
3. 上のデータで、最低点は何点であるか。
4. 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

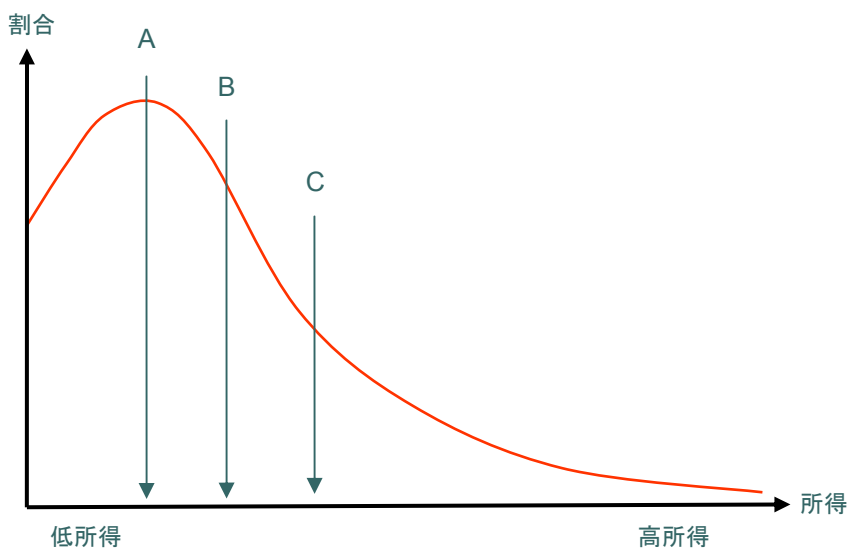


メジアンは中央値であるから、5 つのデータであれば上から 3 番目、下から 3 番目のデータとなる。5 つのデータを並べ替え(sort)ると、90,80,65,60,55 という順になる(降順)。よってメジアンは 65 点。

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

- テストの成績が 50 点、65 点、70 点、40 点、75 点であるとき
 - 平均は何点か。
 - メジアンは何点か。
- 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。

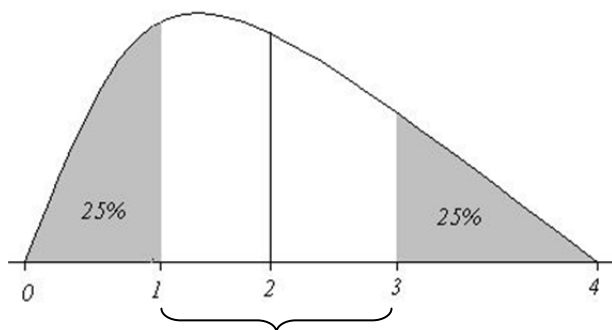
なお、最大値を求める Excel 関数は=MAX、最小値は=MIN である。これはいずれも位置を示している。レンジは、全データ(n 個)のうちで 2 つのデータしか使っていない。すなわち残りのデータ($n-2$ 個)の持つ情報は使われていない。レンジの問題点としては、**外れ値**^{はず}(一つだけとんでもなく異なった値のデータ)の影響を受けやすいことが挙げられる。データ数が大きくなるに従い、レンジも大きくなる傾向があるので注意する。

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①～⑥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。

【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。

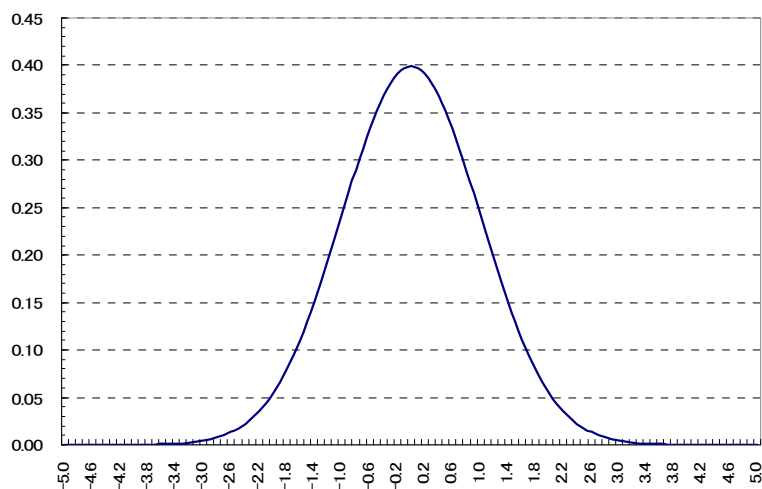
【解説】

□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(density function)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】

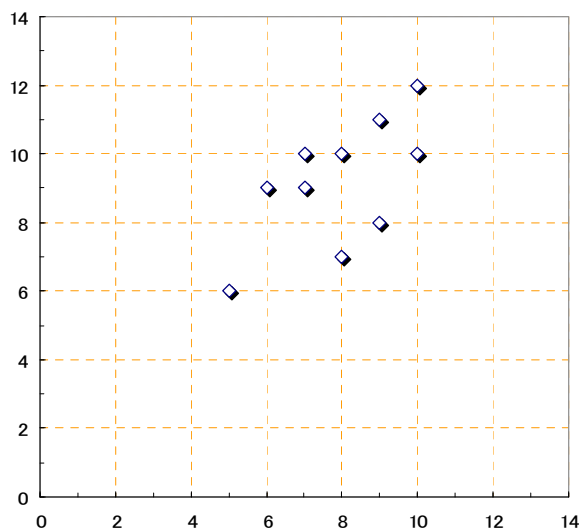
□解説ビデオクリップ



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)



以下は相関係数 (correlation coefficient) の定義式である。

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

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。

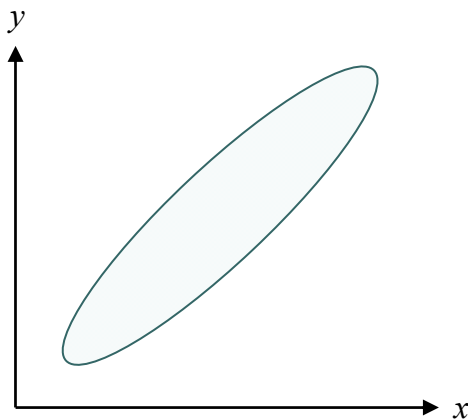
$r > 0$: 正の相関

$r < 0$: 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。

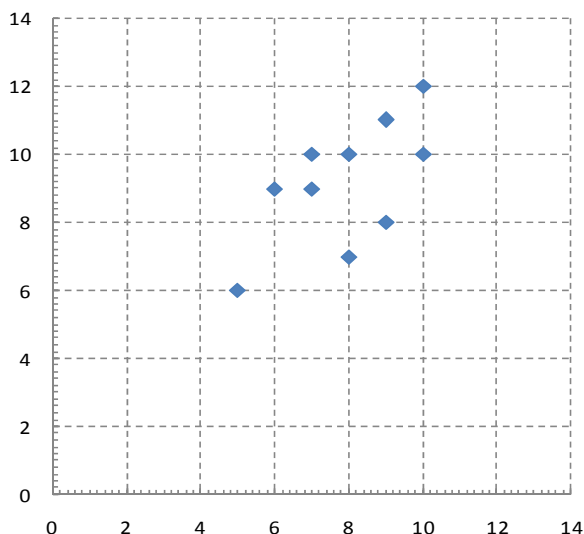
【関連問題】

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



上の表にある 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。例えば、 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

- 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。
- $Y = a + bX$ 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



データの中心を通ると思われる直線と値との差を残差 (residual)あるいは、誤差 (error、略語は e)と呼ぶ。データを現実の値とすれば、直線上にある点は推定された値となる。ここでは右図のように縦方向(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。

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

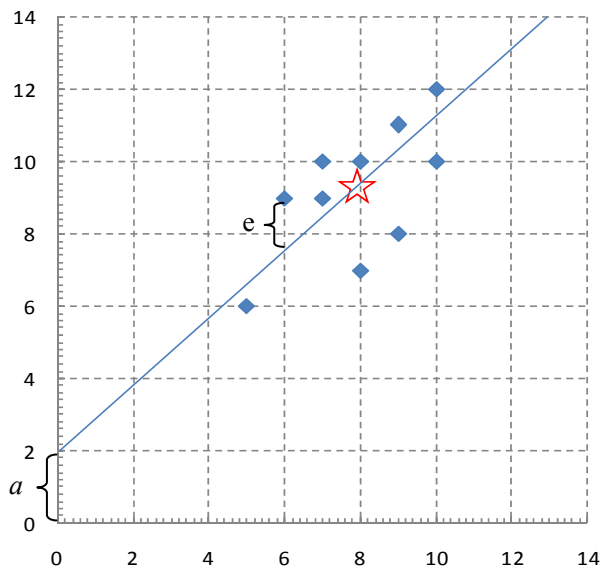
推定誤差 e を 0 に近づけることは、推定誤差が小さくなることなので、中心を通る直線の基準として望ましい。そこで、各推定誤差の和 ($\sum e$) は、以下のような S で表すことができる。

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

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。符号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 a, b を求める公式は以下の通りである。

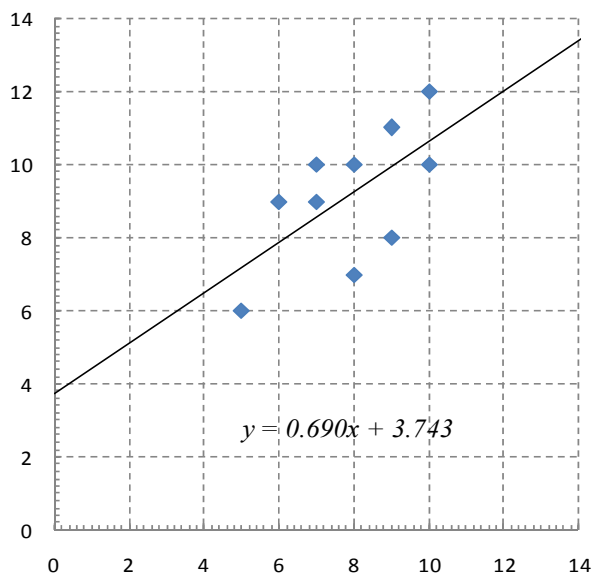
$$a = \bar{Y} - b\bar{X}$$

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

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(偏微分)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。



👉 #063 偏微分

【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

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

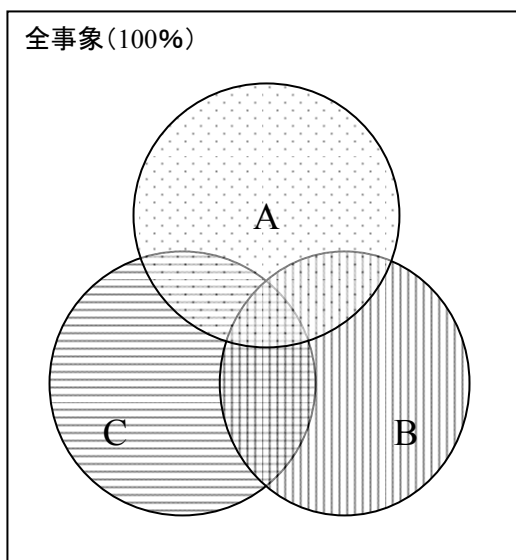
3. 少なくとも 1 回、ヒットを打つ確率

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

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。



📖 #139 リスクと預金保険

【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)から得られる効用水準を計算し、データとグラフを完成させる。

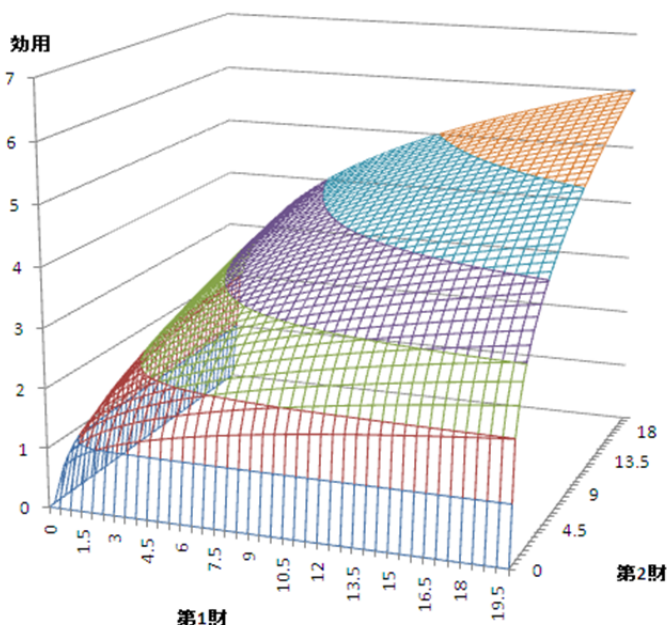
【解説】

□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、右図のような滑らかな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、効用曲面を水平に切断した時の切り口が無差別曲線(indifference curve)である。



☞ #031 無差別曲線

☞ #037 生産関数

☞ #063 偏微分

【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。

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。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。

👉 #038 諸費用の概念

【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(1970 年～2011 年)の月次データ(終値)の推移を表している。グラフの縦軸にメモリアル(値と単位)を記入する。



【解説】

□解説ビデオクリップ



上のグラフは、株価指標の一つである日経平均株価を利用している。横軸は年月を示しており、左端は1970年1月、右端は2011年12月である。日経平均株価が最も高かったのは、1989年12月29日の38,915円である。したがって、縦軸の一番上は45,000円、一番下は0円である。目盛線は5,000円刻みである。**ブラックマンデー**は史上最大規模の世界的株価暴落で、日経平均株価が最高値になる2年前(1987年10月19日の月曜日)に起こっている。

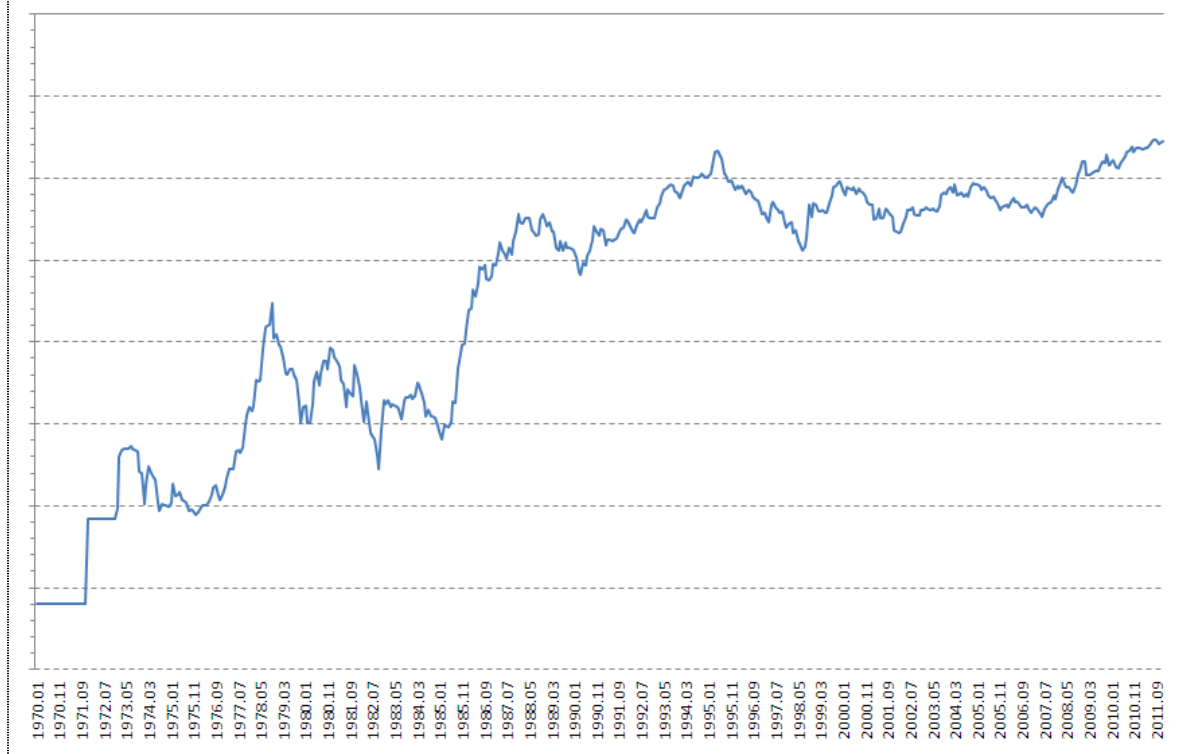
📌 #140 株価指標

【関連問題】

年 月 日

1. 「300 題」のサイトから株価のデータをダウンロードし、Excel で上のグラフを描く。
2. ブラックマンデーはグラフのどのあたりか。
3. 日経平均株価が最も高かったのはいくらか。
4. リーマン・ショックはグラフのどのあたりか。

以下のグラフは、1970 年～2011 年の円ドルに対する為替レート(月末データ)の推移を表している。グラフの縦軸にメモリアル(値と単位)を記入する。



【解説】

□解説ビデオクリップ



上のグラフは、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。横軸は年月を示しており、左端は1970年1月、右端は2011年12月である。縦軸の一番上は0円、一番下は400円で、目盛線は50円刻みである。通常のグラフとは逆(Excelでの軸の反転)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。1972年2月まで固定相場制だったことは、左端が一定値になっていることからわかる。

👉 #153 ニクソンショック

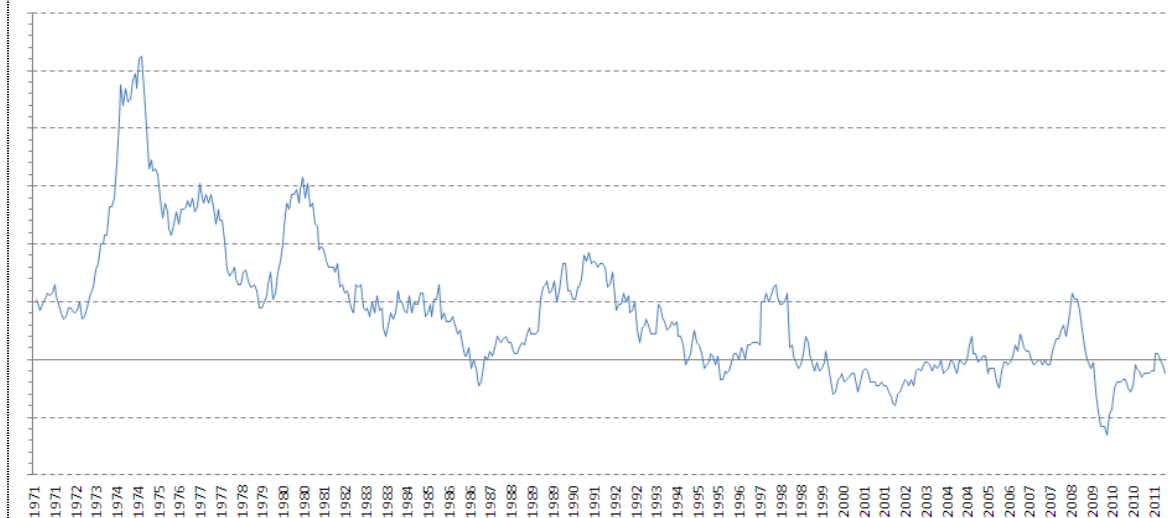
👉 #157 為替介入

【関連問題】

年 月 日

1. 「300 題」のサイトから為替のデータをダウンロードし、Excel で上のグラフを描く。
2. 1972 年では1ドル何円であったか。
3. ブラザ合意はグラフのどのあたりか。

以下のグラフは、1971 年～2011 年の消費者物価指数(対前年同月比)の推移を表している。グラフの縦軸にメモリアル(値と単位)を記入する。



【解説】

□解説ビデオクリップ



上図は、総務省統計局から消費者物価指数(Comsumer Price Index)をダウンロードし、前年同月の差を計算し、その値をグラフ化したものである。横軸は年月を示しており、左端は1971年1月、右端は2011年12月である(前年比の計算から1970年の12ヶ月分のデータはグラフに反映されない)。縦軸にある目盛線で実線が0%であり、一番上が12%で下は-4%である。目盛線は2%刻みである。0%以上である時にはインフレ、下回っている時にはデフレ傾向であることがわかる。2度にわたるオイルショックでは、CPIが急激に上昇している。

👉 #030 フィリップス曲線と自然失業率仮説

👉 #138 物価とスタグフレーション

【関連問題】

年 月 日

1. 「300 題」のサイトから物価のデータをダウンロードし、Excel で上のグラフを描く。
2. 消費税が導入された時期はグラフのどのあたりか。
3. リーマン・ショックはグラフのどのあたりか。
4. 物価のデータとして CPI の他にどのような指標があるか。

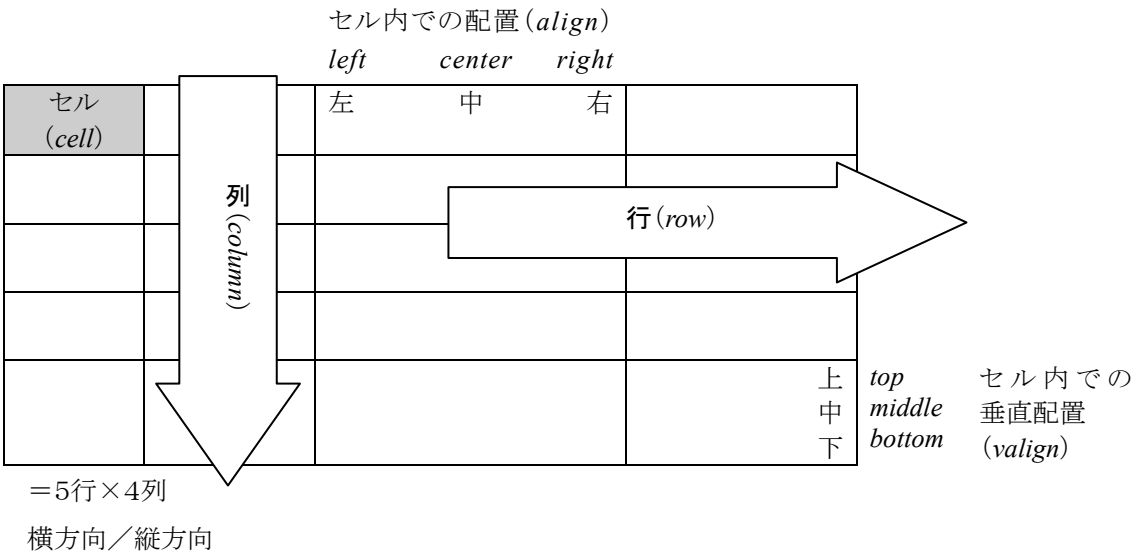
Excel では表のようなシートで計算処理ができる。Excel2010 の1枚のシートには、全部でいくつのセルがあるか。それは 2 の何乗か。

【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。Excel2010 の1枚のシートの中に**行** (row) は $1,048,576 (=2^{20})$ 、**列** (column) は $16,384 (=2^{14})$ があるので、**セル** (cell) は $1,048,576 \times 16,384 = 17,179,869,184 (=2^{20} \times 2^{14} = 2^{20+14} = 2^{34})$ となる。Excel のシートには行番号 (数字) と列番号 (英字: A～XFD) 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。指数表現では 4.57.E+14 である。

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。10,000 は 1×10^4 であるから 1.00E+04 というように表される。E (exponent) の後につく符合が + は桁数が大きいケースで、- は小数点以下を表現する。0.01 は 1×10^{-2} であるから 1.E-02 というように表される。例えば、宝くじ Loto6 の一等の当選確率は $1/43C_6$ であるので、約 0.000016403% である。これを指数表示にすると 1.6403E-07 となる。

スタイル名	入力データ	セル表現
小数点以下 (2 桁)	=pi()	3.14
小数点以下 (14 桁)	=pi()	3.14159265358979
指数	=1/COMBIN(43,6)	1.6403E-07
パーセント小数点以下 (9 桁)	=1/COMBIN(43,6)	0.000016403%

Excel を扱う際の注意として、セルに「10E5」と入力すると「100000」と認識して「1.00E+05」と表示されてしまう。また、「10/1」や「10-1」と入力すると日付データと判断し、「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、0.25 は午前 6 時になる。例えば、名古屋で金環日食が最大となる日時 (2012/5/21 7:31 AM) をシリアル値にすれば「41050.31361」となる。

【関連問題】

年 月 日

- 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。