

平均値・標準偏差値

\$1. はじめに

平均値と標準偏差とだけを求めただけで良いのでしょうか。

これは非常に危険です。

例えば、A社は従業員50人の会社、B社は従業員20人の会社とします。

それぞれの会社の従業員の貯蓄高を調査したところ、

A社：

10万円	2人
20万円	4人
30万円	6人
40万円	8人
50万円	10人
60万円	8人
70万円	6人
80万円	4人
90万円	2人

B社：

50万円	6人
94万7210円	2人
5万2790円	2人

であったとします。

A社、B社ともに、貯蓄高の平均の値は50万円、その分散の値 (式の分母は N) は400万円となります。この結果により、A社とB社とは、その貯蓄高の統計解析からみて、同じと判断して良いのでしょうか。この辺が、本ソフト開発の出発点です。

\$2. 使用法

1. マイクロソフト Excel を起動します。

(本アプリケーションは、マイクロソフト Excel を必要とします。)

2. マイクロソフト Excel のセルに、データを入力します。



入力するデータは Excel のセルに、次の例のように、ランダム (無作為) に置いたとしても OK です。

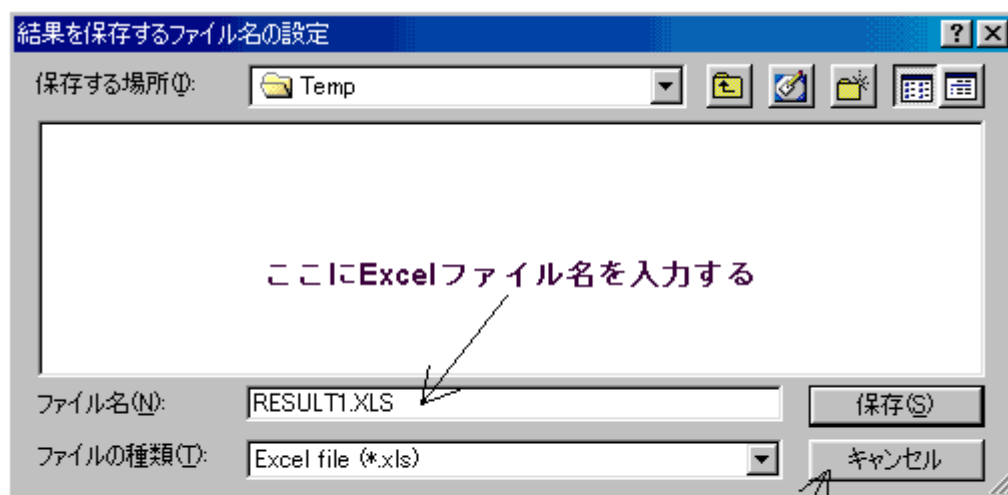
	A	B	C	D	E
1	1	2			
2	2.5		3		
3		3.3	3.6	3.9	
4	4.1	4.3	4.5		
5		4.7		4.9	
6	5	5.1	5.2	5.3	
7	5.5		5.9	6	
8		6.5	6.9		
9	7.2	7.7		8.1	
10	9				
11					

3. データを入力した左上角のセルでマウスの左ボタンを押し、左ボタンを押したままデータを入力した右下角のセルまでマウスをドラッグし、そこでマウスの左ボタンを離します。

例を下記に示します;

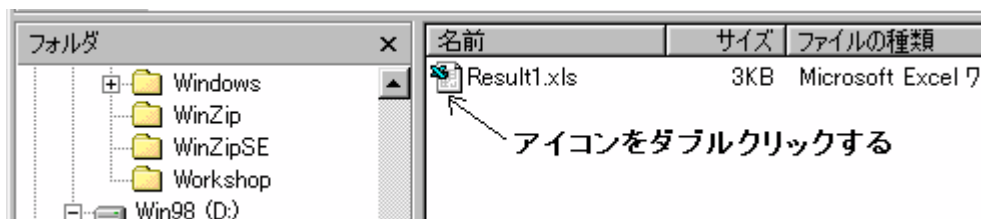
	A	B	C	D
1	1	2		
2	2.5		3	
3		3.3	3.6	3.9
4	4.1	4.3	4.5	
5		4.7		4.9
6	5	5.1	5.2	5.3
7	5.5		5.9	6
8		6.5	6.9	
9	7.2	7.7		8.1
10	9			

4. Excel のメニューの <編集> → <コピー> 選択することによって、データをクリップボードにコピーします。
注: 入力したデータを Excel ファイルとして保存する場合は、まず #2 のステップの時点で データを保存し、その後 #3~#4 のステップを実行してデータをクリップボードにコピーします。
5. マイクロソフト Excel を閉じます。
6. 以上の準備のもとで、本アプリケーションを立ち上げます。
7. 必要なら、<条件設定>ボタン () を押し、解析条件を設定します。
8. <データ解析>ボタン () を押します。
9. エラーがなければ、得られた結果を保存するダイアログボックスが表示されますので、Excel のファイル名 (xxxx.XLS と云うような拡張子が XLS のファイル名) を入力し、<保存>ボタンを押します。



キャンセルを押しても、計算は実行されるので、グラフの表示は可能です

10. 得られた結果の保存後、成功した旨のメッセージが表示されます。
11. 保存した Excel ファイルを、適当なファイラー (例えば、エクスプローラ) で探し、このファイルをマウスの左ボタンでダブルクリックすると、マイクロソフト Excel が起動し、得られた結果を読むことができます。



注: 起動した Excel のファイルの表の一部が "#####" のようになることがあります。
この場合には、下記の手順を実行してください。

- (1) マウスをデータ領域の左上角から右下角までドラッグし、データ領域の背景色を反転させます。

(2) Excel の<書式> → <セル...>を選び、ダイアログの <表示形式>タブ を選択します。

(3)ダイアログの選択ボックスで "標準" を選び、<OK>ボタン を押します。

\$3 . Excel のセルへのデータ入力

1. Excel のセルが空 (つまり、データが無い) ことを確認し、データを入力するセル配列の左上角のセルでマウスの左ボタンを押し、左ボタンを押したまま、データを入力するセル配列の右下角のセルまでマウスをドラッグし、そこでマウスの左ボタンを離します。

2. Excel の<書式> → <セル...>を選び、ダイアログの<表示形式>タブを選択します。

3.下記の手順のどちらかを行います。

(1)通常の数値形式での数値入力の場合;

ダイアログの選択ボックスで"標準"を選び、<OK>ボタンを押します。

この場合、下記の#4 のステップで、通常の数値形式での数値入力 (例えば、10, -3, 14.32, -0.03333, 等) ができます。

入力が許される文字は、

0 1 2 3 4 5 6 7 8 9 . -

です (注:すべて英数半角文字です)。

(2)文字列形式での数値入力の場合;

ダイアログの選択ボックスで "文字列" を選び、<OK>ボタン を押します。

この場合、下記の #4 のステップで、文字列形式での数値入力 (例えば、12, -3.5, 1.3E-2, -2/3, 3/7E-2, 3E-2/7, 等) ができます。

注: 3/7E-2 は 3/(7E-2)、3E-2/7 は (3E-2)/7 を意味します。

入力が許される文字は、

0 1 2 3 4 5 6 7 8 9 . - E /

です (注:すべて英数半角文字です)。

例を下記に示します;

	A	B	C	D
1	1/3	2		
2	2.5		3.5E2/9.1	
3		3.3	3.6	3.9
4	4.1	4.3E-1	4.5	
5		4.7		4.9

4. この準備のもとで、Excel のセルへのデータ入力を行います。

\$4 . その他

1. ヒストグラム

データの数 を N とすると、表示する棒グラフの棒の数は

$$3.3 \cdot \log_{10}(N)$$

が最も良いと云われています。

2. 信頼性の区間推定

平均値、標準偏差、それぞれの信頼性の区間推定では、母集団は正規分布と仮定しています。

3. 保証

このソフトウェアは、フリーウェアで、著作権は下記の著作者が所有しています。なお、本ソフトウェアの使用によって生じる如何なる損害に対しても著作者はその責を負いませんので御了解ください。

\$5. 補足

m: 平均(不偏推定値)、 V^2 : 分散(不偏推定値)、s: 標準偏差(不偏推定値) として

$$m = \frac{\sum_{i=1}^N X_i}{N}, \quad V^2 = \frac{\sum_{i=1}^N (X_i - m)^2}{N-1}, \quad s = \sqrt{\frac{N-1}{2}} \frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \sqrt{V^2}$$

ここで、 $\Gamma(x)$ は ガンマ関数。

© 神田 公生