

「度数分布表」と「ヒストグラム」の作成

ここに「受験した入学試験の成績データ」や「20歳台日本人の身長や体重のデータ」があったとする。そのような場合、そのデータ分布がどの様になっているのか、自分がどの程度のレベルに属するのか知りたくなる。このような場合、「度数分布表」や「ヒストグラム」を作成するのも一つの手であろう。

ここで、データの頻度(重複度数)を数えて表にしたものが「度数分布表」であり、それをグラフに描いたものが「ヒストグラム」または「度数分布図」と呼ばれるものである。なお、数値だけでなく、文字(たとえば、ある英語文書におけるアルファベット)の出現頻度数も同様に扱うことができる。

以下に、そのための手法として、次の3種類の手法について例題を通して学習することにする。

- (1) Countif 関数の利用
- (2) Frequency 関数の利用
- (3) 分析ツール(ヒストグラム)の利用

この中で(3)の手法が一番手軽である。ただし、これがツールメニューに表示されない場合は、分析ツールを組み込んでおく必要がある。すなわち、メニューバー ツール アドイン 分析ツールにチェック✓を入れる。

《問1》下左表は、あるクラスで試験を行なった時の採点結果である。但しクラスは20人で、試験の配点は8点満点で、全て整数で採点されているものとする。度数分布表やヒストグラムを、上の3種類の手法を利用して下図に示すような結果を出せ。

A B C D E F G

20人のクラスの試験の採点結果(試験の配点は8点満点、全て整数で採点)

生徒番号	採点結果
1	5
2	5
3	4
4	7
5	7
6	8
7	5
8	7
9	6
10	8
11	5
12	4
13	3
14	0
15	5
16	8
17	7
18	0
19	4
20	5

得点階級	countifによる度数	frequencyによる累積度数	frequencyによる度数の算定
0	2	2	2
1	0	2	0
2	0	2	0
3	1	3	1
4	3	6	3
5	6	12	6
6	1	13	1
7	4	17	4
8	3	20	3

得点階級	得点階級	頻度
0	0	2
1	1	0
2	2	0
3	3	1
4	4	3
5	5	6
6	6	1
7	7	4
8	8	3
	次の級	0

ヒント:

E4 セル=COUNTIF(\$B\$6:\$B\$25,D4) F4 セル=FREQUENCY(\$B\$6:\$B\$25,D4) G5 セル=F5-F4

範囲\$B\$6:\$B\$25内でセルD4の値より小さい値のデータ個数をカウント
範囲\$B\$6:\$B\$25内のデータについてセルD4の値に一致するならば、そのデータ個数をカウントする。

メニューバー ツール 分析ツール ヒストグラム.....

《問2》下左表はあるクラスの成績結果である。但し、クラスは20人で、試験の配点は100点満点で、全て整数で採点されている。度数分布表やヒストグラムを、上の3種類の手法を利用して下図に示すような結果を出せ。

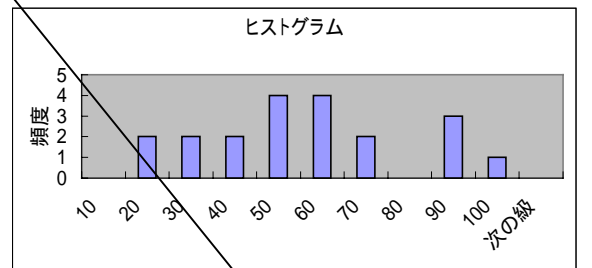
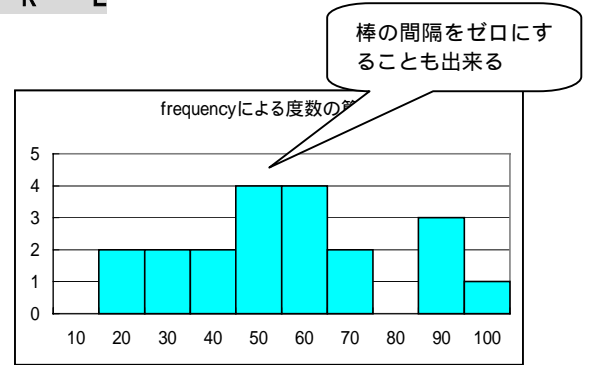
A B C D E F G I J K L

20人のクラスの試験の採点結果(試験の配点は100点満点、全て整数で採点)

生徒番号	採点結果
1	54
2	34
3	65
4	86
5	84
6	95
7	16
8	55
9	82
10	38
11	46
12	26
13	23
14	40
15	45
16	18
17	49
18	52
19	63
20	53

得点階級 (下)	得点階級 (上)	countifによる累積度数	frequencyによる累積度数	frequencyによる度数の算定	手と目による算出の結果確認
0	10	0	0	0	0
10	20	2	2	2	2
20	30	4	4	4	2
30	40	6	6	6	2
40	50	10	10	10	4
50	60	14	14	14	4
60	70	16	16	16	2
70	80	16	16	16	0
80	90	19	19	19	3
90	100	20	20	20	1

得点階級	得点階級	頻度
10	10	0
20	20	2
30	30	2
40	40	2
50	50	4
60	60	4
70	70	2
80	80	0
90	90	3
100	100	1
	次の級	0



ヒント:

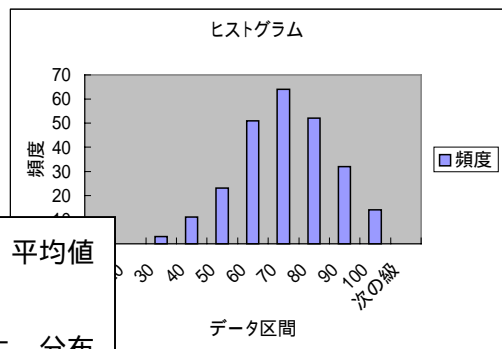
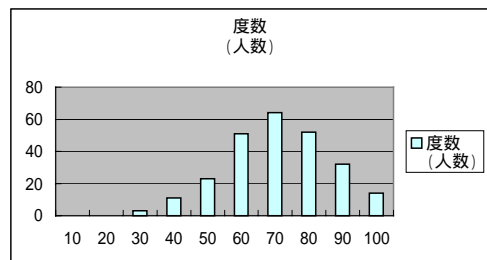
F4セル=COUNTIF(\$B\$6:\$B\$25,"<="&E4) G4セル=FREQUENCY(\$B\$6:\$B\$25,E4) H5セル=G5-G4
 範囲\$B\$6:\$B\$25内でセルE4の値より小さい値のデータ個数をカウント
 範囲\$B\$6:\$B\$25内のデータについてセルE4の値より小さいデータならば、そのデータ個数をカウントする。

《問3》ここに、下表のような250人の成績データがあるとすると、3種類の方法を用いて、度数分布表とヒストグラムを描け。

生徒名	成績
生徒1	60
生徒2	46
生徒3	69
生徒4	84
生徒5	83
生徒6	91
生徒7	32
生徒8	61
生徒9	81
生徒10	49
生徒11	55
生徒12	40
生徒13	37
生徒14	50
生徒15	53
生徒16	33
生徒17	56
生徒18	59
生徒19	67
生徒20	60
生徒21	60
生徒22	59
生徒23	85
生徒24	64
生徒25	62
生徒26	57

範囲下	範囲上	frequency関数を利用	countif関数を利用	度数(人数)
0	10	0	0	0
10	20	0	0	0
20	30	3	3	3
30	40	14	14	11
40	50	37	37	23
50	60	88	88	51
60	70	152	152	64
70	80	204	204	52
80	90	236	236	32
90	100	250	250	14
				250

区間範囲	データ区間	頻度
10	10	0
20	20	0
30	30	3
40	40	11
50	50	23
60	60	51
70	70	64
80	80	52
90	90	32
100	100	14
	次の級	0



《問4》250人分の成績データ(但し100点満点、平均値60、標準偏差10)を作成せよ。

《解》メニューバー ツール データ分析 乱数発生 分布を「正規」,変数の数,乱数の数を「1」「250」.....