

ヒストグラムpp1~8

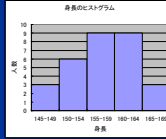
■ 高2女子30人の身長

■ 145,146,148,150,152,
152,153,153,153,155,
155,155,156,157,158,
158,158,159,160,160,
160,160,160,161,161,
162,164,166,166,167

身長 (cm)	人数
145以上150未満	3
150~155	6
155~160	9
160~165	9
165~170	3
計	30

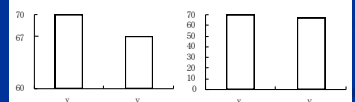
身長ヒストグラム

- 身長は色々な値をとるので変量
- 変量を区間に区切る
- 区間を階級、階級の幅を階級幅
- 各階級に属する要素の数を度数
- 度数分布のグラフをヒストグラム

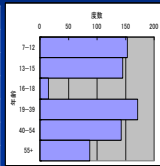


ヒストグラム作成における注意

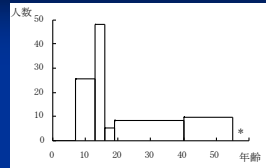
- Xを服用した方のうち70人が治癒
- Yを服用した方のうち67人が治癒



年齢	度数	階級幅	度数/幅
7-12	153	6	25.5
13-15	145	3	48.3
16-18	15	3	5.0
19-39	171	21	8.1
40-54	142	15	9.5
55+	87	*	*

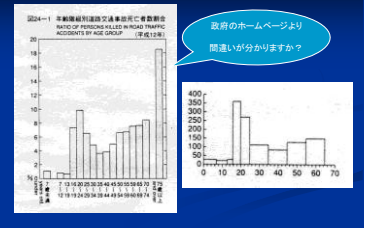


柱の太さを階級幅にする



度数分布表を作成するときは、階級幅を一定にするのが賢明

柱の太さを階級幅にする



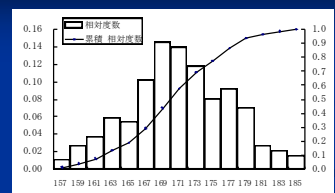
高2男子186名の身長の分布

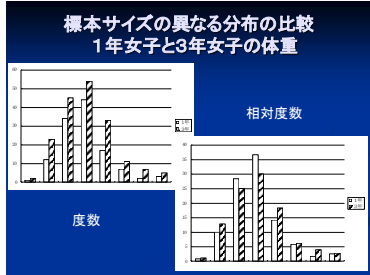
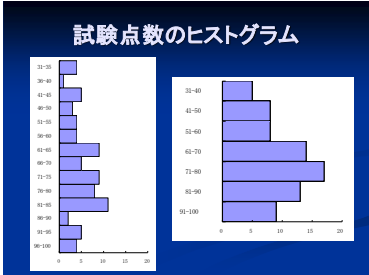
身長	度数	相対度数	累積相対度数
155-157未満	2	0.011	0.011
157-159 "	5	0.027	0.038
159-161 "	7	0.038	0.075
161-163 "	11	0.059	0.134
163-165 "	10	0.054	0.188
165-167 "	19	0.102	0.290
167-169 "	27	0.145	0.435
169-171 "	26	0.140	0.575
171-173 "	22	0.118	0.694
173-175 "	15	0.081	0.774
175-177 "	17	0.091	0.866
177-179 "	13	0.070	0.935
179-181 "	5	0.027	0.962
181-183 "	4	0.022	0.984
183-185 "	3	0.016	1.000

相対度数

- 相対度数=度数/全数
- 相対度数の和は1
- 相対度数表を身長の分布表とよぶ
- 相対度数の累積和を累積相対度数

身長ヒストグラム





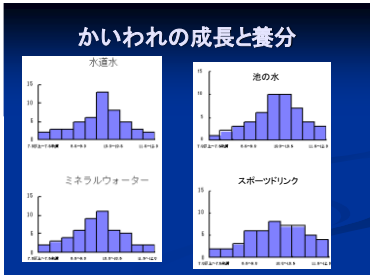
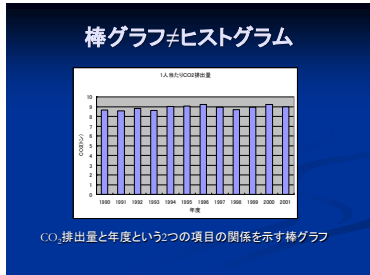
ヒストグラムの歴史

- グロント
 - ベストによる死亡票を収集し年齢5歳刻みの死亡率を求めた
- 死亡率は年齢で著しく異なるが、5歳刻みに分類すれば、級内での死亡率は著しくは異なる!
- 収集した資料を似た特性を持つものに分類し、各群の特徴を比較する方法

ヒストグラムの見方

まず全体の形状を眺める

1. 単峰か
2. 左右対称か、非対称か
3. 凸凹しているか
4. 両端まで自然に減少
5. 外れ値はあるか
6. 階級幅が異なるか
7. 原点を含むか



外れ値の処理

- 集団から離れて存在する値
 - そのときだけ実験条件が狂っていた?
 - あるいは単なる転記ミス?
- 原因が分かれば適切な処理が可能
 - 外しても含めても結論が同じなら、気にしない
 - 結論が変わる場合は、問題!
- 外れ値の目安
 - 75%点 + 1.5(75%点 - 25%点) を越える値は外れ値とする
 - 25%点は下から1/4の値
 - 75%点は上から1/4の値

オゾンホールが発見と外れ値

- 1985年に英国の科学者が南極上空オゾン層にホールを発見したが、この発見は無視された
 - 地上から観測していたので信頼できない
 - 人工衛星から見下ろしたデータからホールは無いとされた!
 - 見落とした原因は?
- 膨大なデータを処理するソフトに、低いオゾン値は外れ値とする機能
- 1979年の人工衛星からの観測データを再分析してホールの存在を確認