ヒストグラム範囲の決め方

[統計]

・ <http://www.ipc.shimane-u.ac.jp/food/kobayasi/how%20to%20write%20a%20histogram.htm>

範囲の決め方

1. データの数 n を求める。度数分布表を作るには n は少なくとも 50、できれば 100 以上必要である。
2. データを大きさの順に並べる。
3. 範囲 R を求める。データの中から最大値 $\max$ と最小値 $\min$ をみつけ、範囲を計算する。
 $R = \max - \min$
4. 階級間隔 h を決める。範囲 R を 10 ~ 20 の適当な数で割る（この商を k とする）。階級間隔を測定値の測定単位の整数倍とするために k を整数値になるように丸める。このときデータ数が 100 より少ないときは k を大きめの値に丸め、100 以上のときは k を小さめの値に丸める。
5. 階級分けの出発点を決める。最初の階級は最小値を含み、階級間隔 h であり、次の階級との境界値は共通し、しかもデータが 2 階級にまたがらないように、決めなければならない。そのため、最小値から測定単位の 2 分の 1 を引いた値を階級の出発点とするとよい。（こうすると階級の境界値は測定単位の整数倍プラス測定単位の 2 分の 1 となり、境界値にデータが重なることはなくなる。）
6. 出発点の値に階級間隔 h を逐次加え、最大値を含む階級までの境界値を決定する。
7. 度数分布表を作成する。各階級の中央の値をその階級の階級値とする。
8. 次にヒストグラムを作成する。横軸に階級の境界値を取る。
9. 縦軸に度数を目盛る。このとき、できる図形の縦・横の比が 2:3 程度になるのがよい。