

2004.4.27 現代日本論演習 I (田中重人)

第3回「統計分析の基礎」目次

1. データ収集から分析まで
2. 変数の種類
3. 度数分布表とヒストグラム

【データ収集の方法】

●実験とは

●観察とは

※ 「調査」は、通常、「観察」にふくめる

※ この授業では、実験の方法はあつかわない

【データ分析のふたつの水準】

● データの特徴を少数の数値に要約して記述 = **記述統計**

● 誤差の評価

(この手続きの一部が**推測統計**)

(教科書 p. 1-6)

- データ収集の際は、誤差を
できるだけ少なくする方法を考えること
- 分析の際は、まず手元のデータの
特徴をしっかりとつかむ
→ 記述統計 (教科書 第5章まで)
- そのうえで、データがふくむ「誤差」を
推測し、「真の値」について考察する
→ 推測統計 (教科書 第6章以降) + α

【変数の種類】

- 比率尺度 (ratio scale)
- 間隔尺度 (interval —)
- 順序尺度 (ordinal —)
- 名義尺度 (nominal —)

(質的変数とも)

(教科書 p. 8)

【尺度の変換】

- ★ 上位の尺度のほうがあつかえる演算が豊富
- ★ 上位の尺度は下位の尺度の特徴を兼ね備えている

→分析手法の選択幅がひろい

私たちが測定するものはたいてい
順序尺度以下である

- ★ 上位の尺度への変換には
一定の理論的根拠が必要

【度数分布表】

Frequencies コマンドを使う

- ★ 度数
- ★ 相対度数 (%)
- ★ 累積度数・累積相対度数
- ★ 欠損値のあつかい

(教科書 p. 27–31)

【累積%とパーセンタイル】

- 順序尺度以上の場合のみ意味を持つ
- Percentile(= %点)
- 中央値 (median) = 50%点
- 「割り切れてしまおう」場合は中点をとる
(教科書 p. 43)
- 同じ値が並ぶ場合は多少の操作が必要
(森敏昭・吉田寿夫(編) (1990) 『心理学のための
データ解析テクニカルブック』 北大路書房. p. 15)

【棒グラフとヒストグラム】

- 棒グラフ……棒同士の間空白をあける。高さ(長さ)をよむ。
- histogram (柱グラフ)……柱の間隔をあけない。面積をよむ。

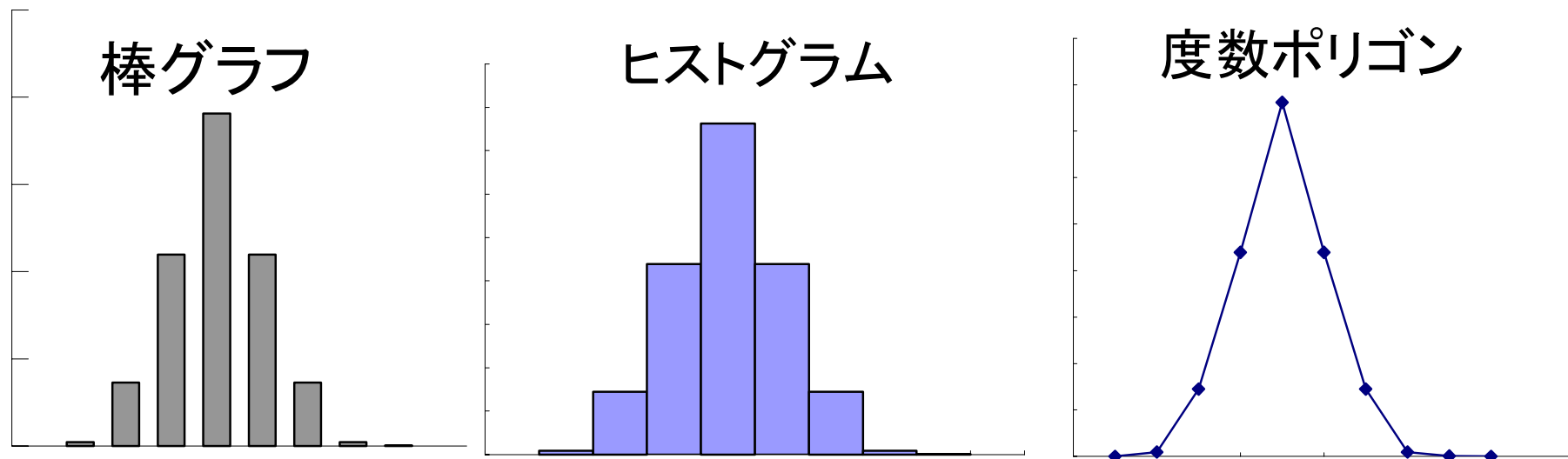
※縦軸は度数または%

★ 連続量を階級分けした場合
→ ヒストグラム

★ それ以外の場合 (離散量/
名義尺度) → 棒グラフ

※度数多角形 (polygon) は複数の変数の分布を比較するときに便利。

(教科書 p. 32-36)



SPSS では histogram が書きにくい。

★ recode で整形した上で度数分布表のメニューで「図表…」指定。棒グラフを書く

★ グラフ→インタラクティブ→ヒストグラム
では等間隔の区間に分割してくれる

【実習】

- (1) 本人年齢の度数分布表を出力し、中央値と上側 20% 点に印をつけよ
- (2) 適当な変数について棒グラフまたはヒストグラムを作成