

## 語句問題対策

※白丸はそれほど重要ではない（気がする）ものをあらわす。

- ある集団について調べるとき、実際にその集団全体を調べるときを**全数調査**という。一方で、集団の一部について調べ、統計的推測により全体を知ろうとする**標本調査**もある。
- 集めたデータのうち、身長や体重、テストの点数などのように定量的な値で与えられるものを**量的データ**と呼ぶ。一方で、性別や学歴など、状態を表すデータを**質的データ**と呼ぶ。
- 記述統計学**・・・ある集団の特徴を記述するために、観測対象となった各個体について観測し、得られたデータを整理・要約する方法
- シケプリ No.2 または教科書 P18,19 で、**度数分布表**と**ヒストグラム**を理解すること。
- 右に歪んだ分布**は、右から山を押したような形！つまり左に度数が多い。（左はもちろん逆）
- ヒストグラムの山が二か所あるとき、これを**双峰型**という（山が一つなら**単峰型**）。双峰型のとき、これは異なる二つの性質のデータが混じっているためと考えられる。このとき、適当な二つのグループにデータを分けることで、単峰型にできることがある。このグループ分けの作業を**層別**という。
- ローレンツ曲線**・・・二つの変数についての累積度数を平面上にプロットしてつないだ曲線。正方形の対角線からのズレから、二つの変数間の不均衡を判断できる。教科書 P27 で、図を確認すること。
- 平均はともかく、**メディアン・モード・四分位偏差**の定義や意味については押さえておきたい。教科書 P32,33,36 またはシケプリ No.4,5 を見よ。
- 二次元の（量的）データを平面にプロットしたものを**散布図**という。一方で、質的データを含む場合は分割表という。このとき、分割表の縦方向の変数を**表側**、横方向の変数を**表頭**という。

○相関係数について、 $r=0.5$  くらいで、もうあまり相関はない。

○分布の非対称性についての指標に歪度がある。式は教科書 P100 で確認のこと。 $\alpha = 0$  のとき左右対称。 $\alpha > 0$  のとき右の裾が長い。 $\alpha < 0$  のとき左の裾が長い。

○分布の尖り具合の尺度として、尖度がある。式は教科書 P101 で確認のこと。 $\alpha$  が大きいほど尖っており、正規分布は  $\alpha = 3$ 。これと比較することが多い。

●シケプリ P20~23 に紹介したくらいの確率分布の名前は、覚えておくこと。

●標本から母集団について推定するとき、事前に母集団がどんな分布に従っているかわかっている場合をパラメトリックの場合という。逆に、母集団が従う分布がわかっていない場合をノン・パラメトリックの場合という。

●母集団の未知の母数 (=母集団の確率分布を決定するパラメタのこと。母平均など) を推定する際、区間推定のほかに、ある一つの値で推定する方法があり、これを点推定という。この際、最尤法 (さいゆうほう) という手続きがある。

●母平均と標本平均、母分散と標本分散について、その違いをはっきりとわかっておくこと。←これはテストに出すとか言っていた気がする。できない人いないと思うんだけど、、、

まあ、語句問題やる暇があるなら他の計算問題とかやったほうがいいんじゃないですかね。

また、このプリントは別に前バランがあったところを集めたとかではなく、完全に僕の独断と偏見で出そうな語句を選んだものなので、あんまり信用しないこと！基本的には、あまり難しい内容は扱っていないつもりです。