「判別分析を行い話者の性別を判別する」

2021/5/30

学籍番号：2J19F508-7

氏名：川崎咲希

　各データの韻律情報のうち、基本周波数を対象として判別分析を行い、男性・女性の話者を判別できる基準を求める。

**使用データ：**

　分析に使用したデータは以下のとおりである。

学会講演(女)：A01F0055/A01F0067/A01F0122

学会講演(男)：A01M0007/A01M0015/A01M0020

模擬講演(女)：S00F0014/S00F0031/S00F0041

模擬講演(男)：S00M0025/S00M0053/S00M0065

**【結果】**

1) F0平均値とF0標準偏差の二次元散布図を作成した結果を図１に示す。

2) 男性・女性でプロットの色を分けて作成した二次元散布図を図２に示す。男性・女性でプロットの色を分けた結果、男性は左側、女性は右側に集まっており判別が可能であると考えられる。

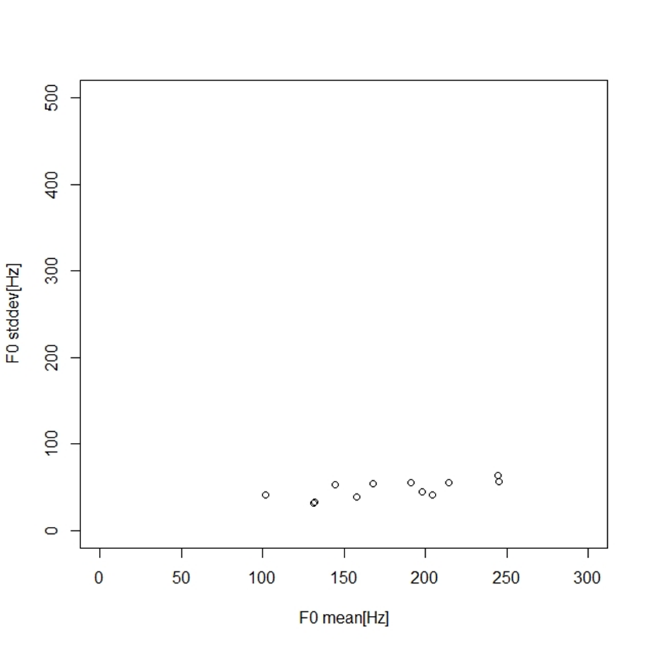


図１　二次元散布図

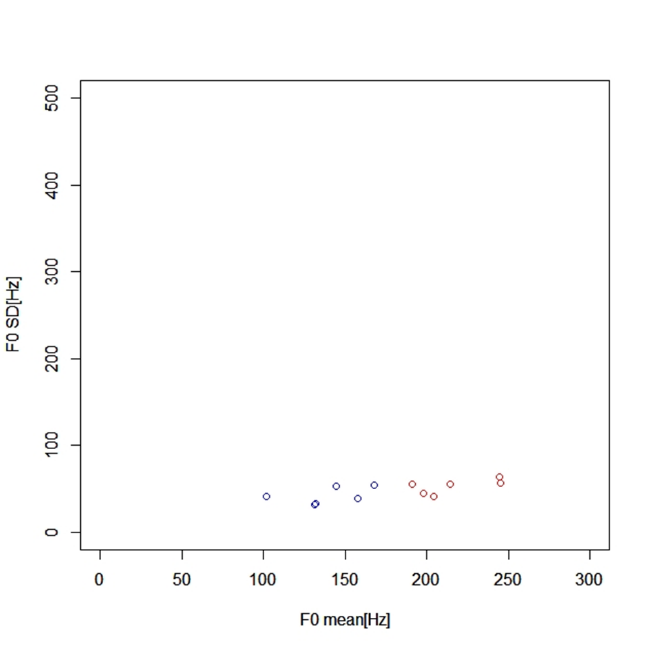


　　図２　二次元散布図（男性：青、女性：赤）

3)判別分析を行った結果、図3のように群平均および線形係数が求められた。以上から次式のような線形判別式が得られる。

Sex = α + (-0.0480193)＊F0平均値 + 0.0290281＊F0標準偏差

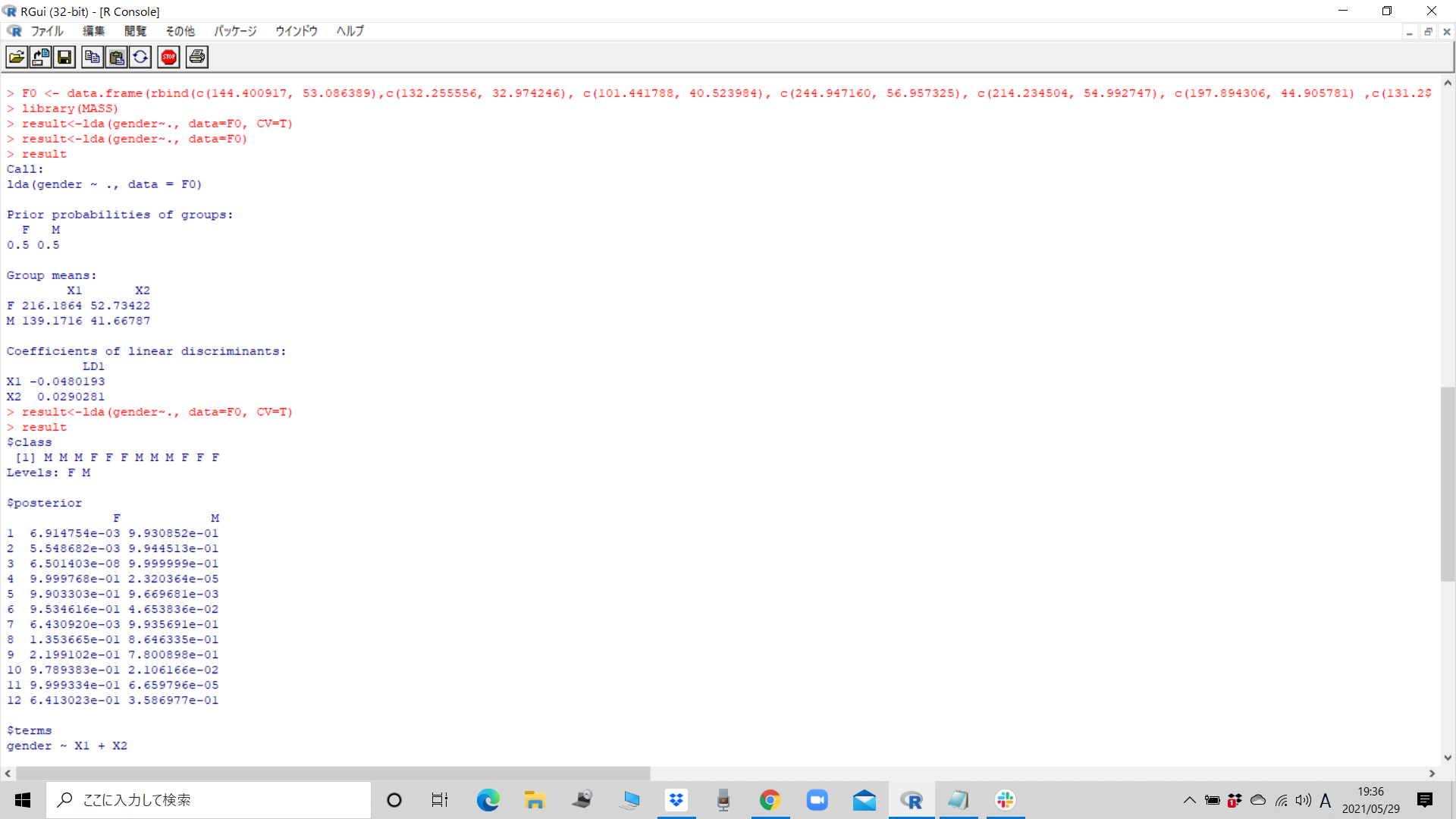


図3　判別分析の結果

4) 交差検定をして判別分析結果の検証を行った結果、図４のように男女の判別がなされた。これは入力したデータの性別と一致していた。

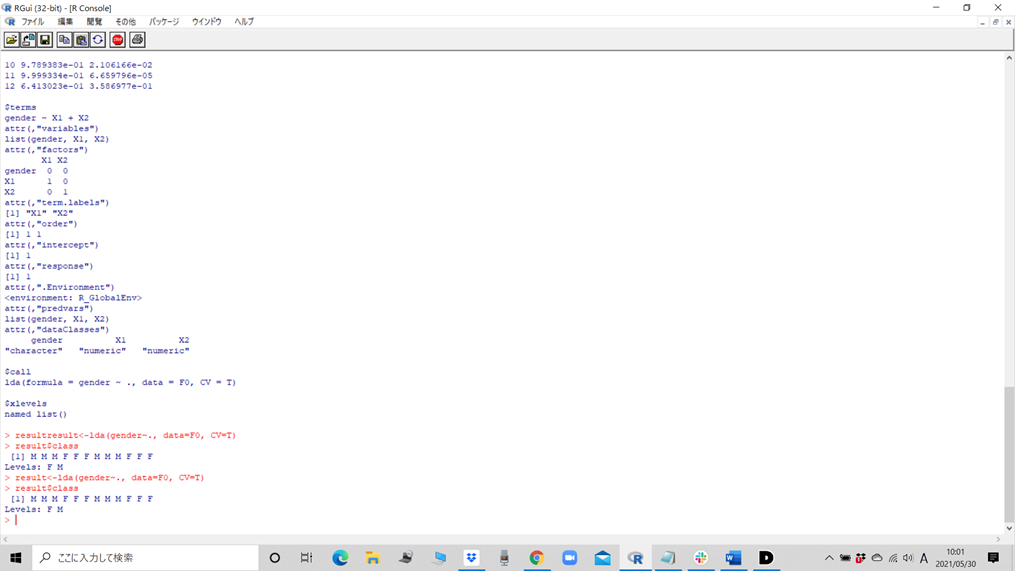


図４　交差検定の結果

**【考察】**

12の話者のデータを用いて判別分析を行い、男女の判別が可能であることがわかった。少ないデータ数で求めた判別基準では、どのくらいの正確さで判別可能なのか興味を持った。

話者の性別の違いを判別するにあたって、基本周波数の値は有用であることを実感した。また、語尾やフィラーなどの性別によって異なる可能性がある要因でも、数値化することができれば判別基準を求められ様々に検証可能であると思われた。