

「生物統計学」2015年度レポート課題

【課題】 データの可視化に関する下記の設問に答えよ。

設問 1) 統計学者 Francis Anscombe が提示した「Anscombe's quartet」(図 1) は、平均や分散あるいは相関係数などの基本統計量だけでなく回帰直線まで完全に一致するにもかかわらず、散布図を描けば全く異なる複数のデータセットの実例である。この図からデータ可視化と統計分析との関係について述べよ。

Cf: Anscombe's quartet

https://en.wikipedia.org/wiki/Anscombe's_quartet

設問 2) データ可視化にはグラフが不可欠である。しかし、グラフがデータの挙動をいつも適切に可視化できるとは限らない。Tracey L. Weissgerber et al. の論文の図 (図 2) から、多用されている棒グラフにはどのような欠陥があるのかを考察せよ。

Cf: Tracey L. Weissgerber et al. 2015. Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. doi:10.1371/journal.pbio.1002128 <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002128>

設問 3) あるデータセットに対して複数のグラフによる可視化が利用できるとき、それぞれのグラフの長所と短所を比較することができる。Michaela Spitzer et al. の図 (図 3) から、棒グラフ・箱ひげ図・バイオリン図・ビーン図を相互比較せよ。

Cf: Michaela Spitzer et al. 2014. BoxPlotR: a web tool for generation of box plots. doi:10.1038/nmeth.2811 <http://www.nature.com/nmeth/journal/v11/n2/full/nmeth.2811.html>

図 1

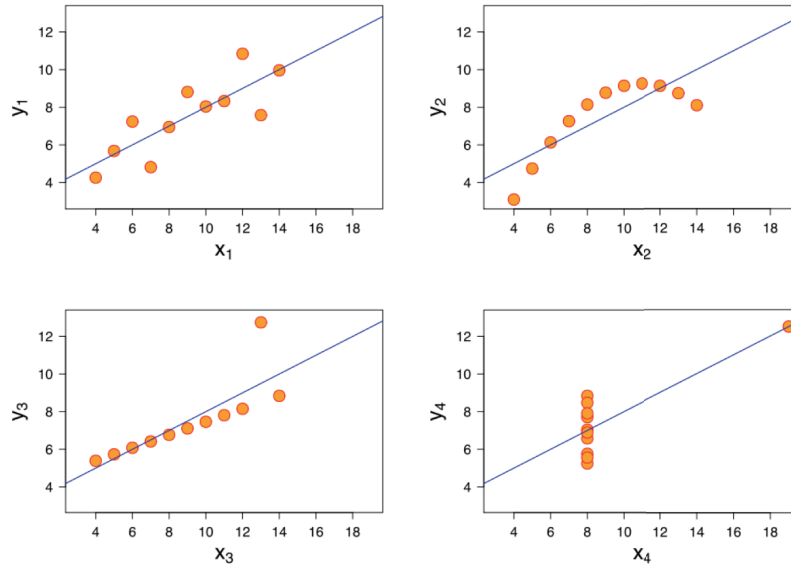


図 2

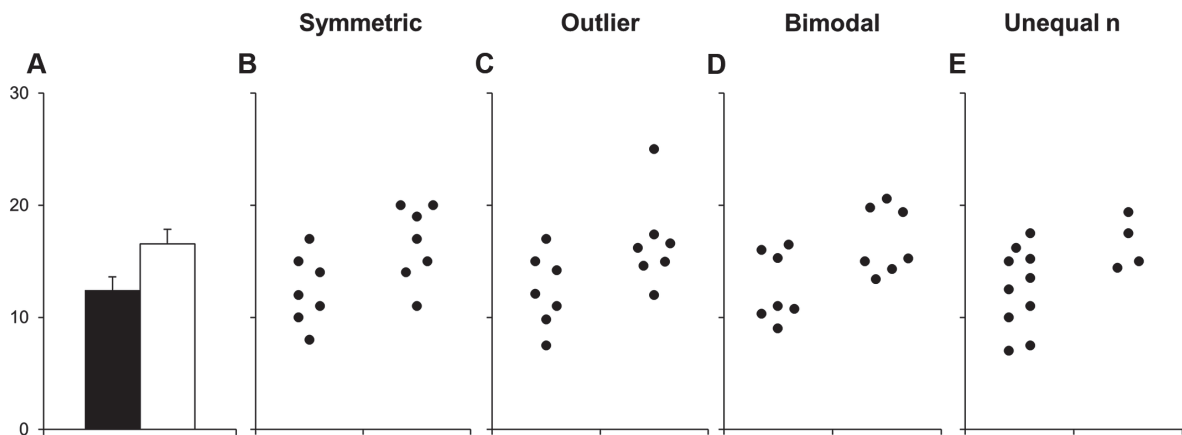
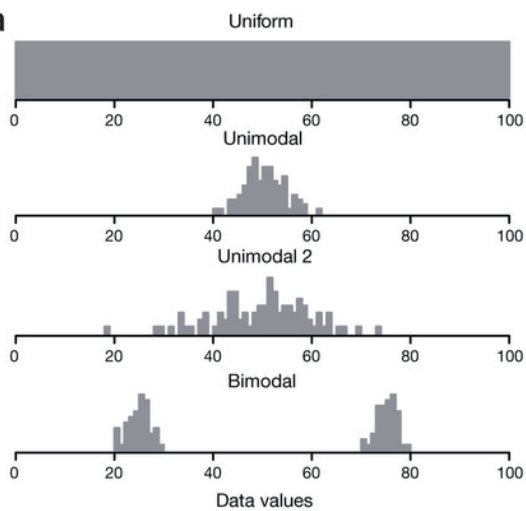
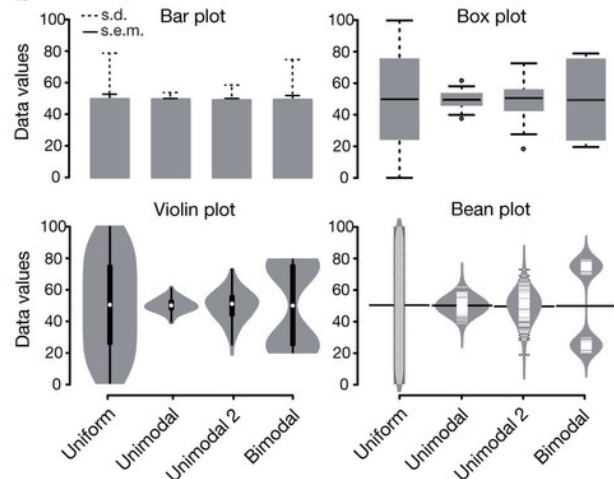


図 3 a



b



【締切日】 2015 年 12 月 17 日 (木)

【提出先】 提出レポートには氏名と学籍番号を明記し，メール本文にプレーンテキストとして書いて，三中信宏 (minaka@affrc.go.jp) 宛てにメール送信すること．一切の添付ファイルは御法度である．また，正常に受信された場合にかぎり受領メールを返信する．送信後一両日中に受領メールが届かなかった場合はレポート不着の可能性があるので再度送ること．