

2013

Breakdowns in Common Statistical and Graphical Techniques for Big Data

John Gabrosek
Grand Valley State University, gabrosej@gvsu.edu

Follow this and additional works at: https://scholarworks.gvsu.edu/bigdata_conference2013

ScholarWorks Citation

Gabrosek, John, "Breakdowns in Common Statistical and Graphical Techniques for Big Data" (2013). *2013 Presentations*. 4.

https://scholarworks.gvsu.edu/bigdata_conference2013/4

This Article is brought to you for free and open access by the Big Data Conference at ScholarWorks@GVSU. It has been accepted for inclusion in 2013 Presentations by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

John Gabrosek (Department of Statistics)

**Breakdowns in Common Statistical
and Graphical Techniques for Big
Data**

Breakdowns in Common Statistical and Graphical Techniques for Big Data

John Gabrosek,
Department of Statistics

Hypothesis Testing

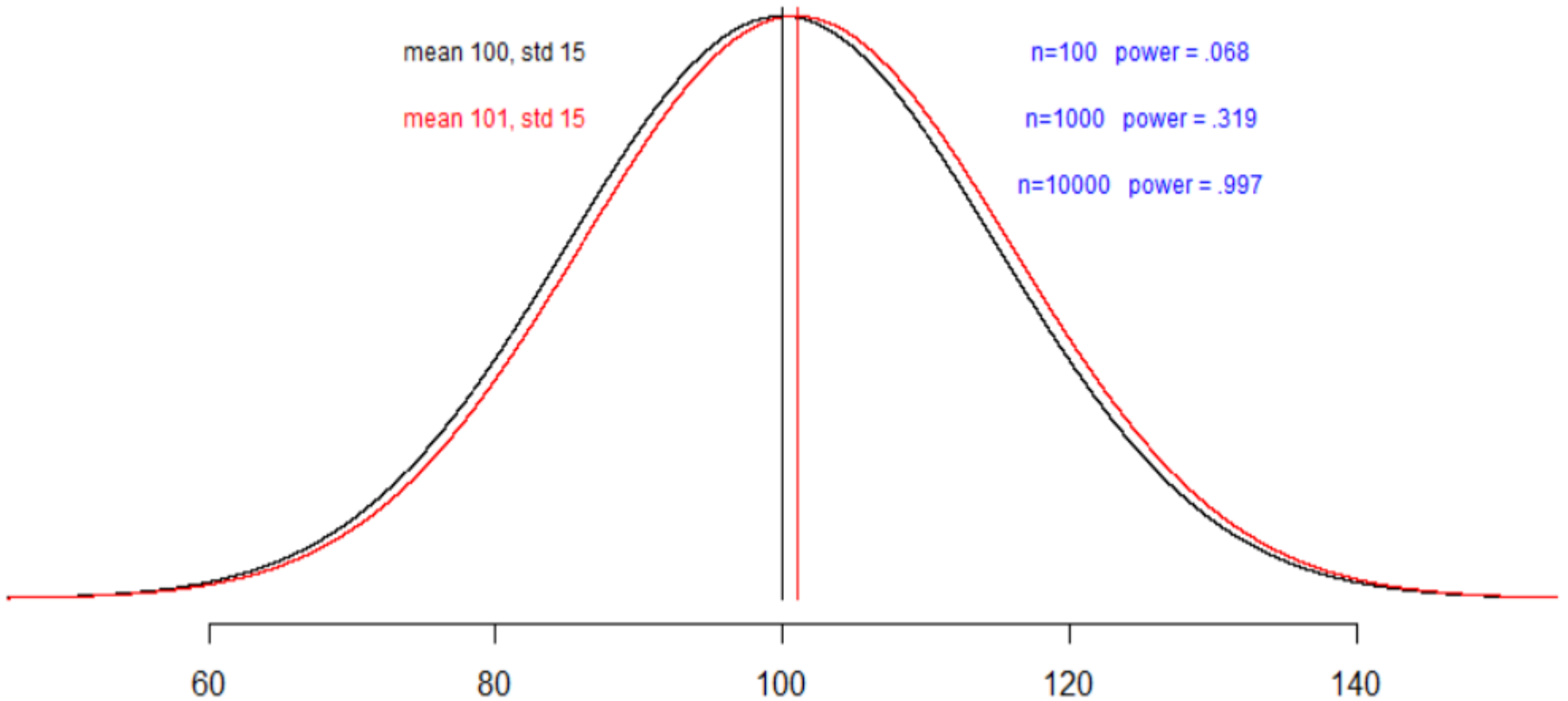
mean 100, std 15

mean 101, std 15

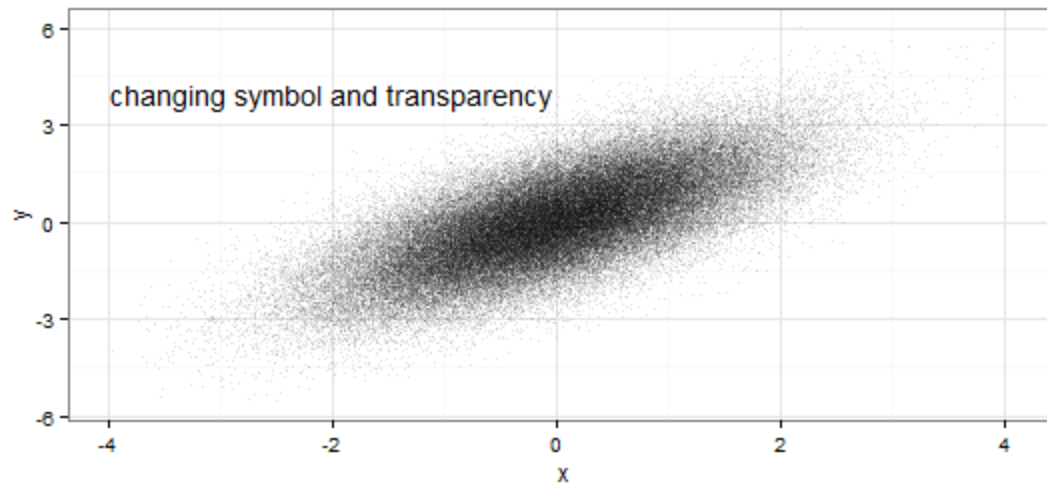
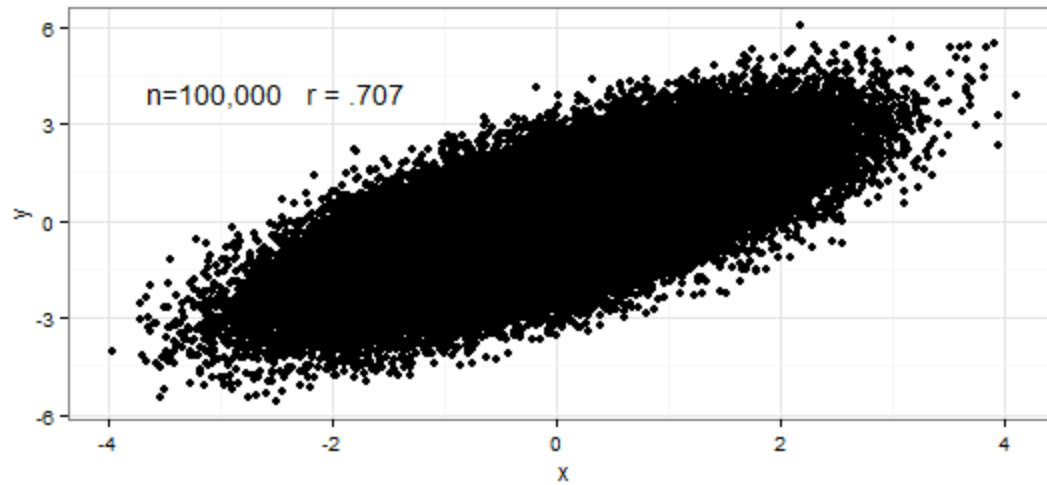
n=100 power = .068

n=1000 power = .319

n=10000 power = .997



Graphics



Other Issues

- Context of the problem
- Confounding and lurking variables
- Sampling and representativeness
- Integration of multiple sources
- Dependent data
- Multivariate data