STAT 6560
Graphical Methods

Spring Semester 2009

Project 1

Brittany Allred

Utah State University

Department of Mathematics and Statistics

3900 Old Main Hill

Logan, UT 84322–3900

Graphics in Statistics Education

Citations: *Justification for graphics in the classroom - Jane Gentleman*

- "*indications are that the computer is on its way into the classroom*" - (Gentleman (1977), p. 166)
- "*Spontaneity is the main advantage of a terminal-equipped classroom.*" - (Gentleman (1977), p. 166)
- "*teaching is more fun for the professor and learning is easier and more fun for the students*" - (Gentleman (1977), p. 167)
- "*[Student's have the] valuable ability to control the details of the examples themselves*" - (Gentleman (1977), p. 171)
- "*Plots, in particular, interactively generated computer plots and, especially, spontaneously produced ones, are extremely useful as an aid to teaching statistics*" - (Gentleman (1977), p. 174)
- "*Statisticians are fortunate that modern computer technology has provided a gold mine of new possibilities. Let us mine the gold, shape it to our own needs, and use it.*" - (Gentleman (1977), p. 174)

Applications in the Classroom - Suggestions by Gentleman

- **Continuous Distributions**
  - Use graphics to compare cdf's and pdf's

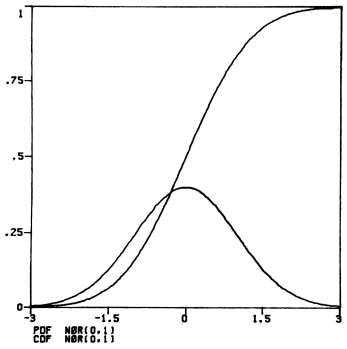

Figure 1: Gentleman (1977), p. 167

- **Discrete Distributions and Approximations**
  - Use graphics as simulations of various distributions and how they do and do not converge.
  - A simulation with binomial data with Poisson distribution superimposed.

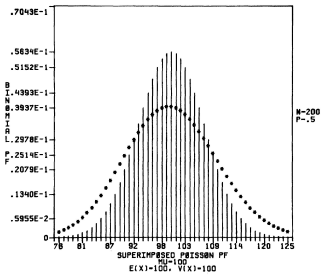

Figure 2: Gentleman (1977), p. 168

- **Histograms and Goodness-of-Fit**
  - Plot data from the class and superimpose distributions over the data to show the students how the data fits a specific distribution.
  - R Example: histograms with real data with the normal distribution superimposed on top. Snell & Peterson (1992), Gentleman (1977)

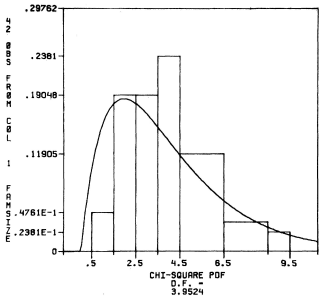

Figure 3: Gentleman (1977), p. 170

- **Central Limit Theorem**
  - Creating histograms with various sample sizes demonstrating the more in the sample, the closer the shape of the histogram will be to the normal distribution. Tijms (1992)

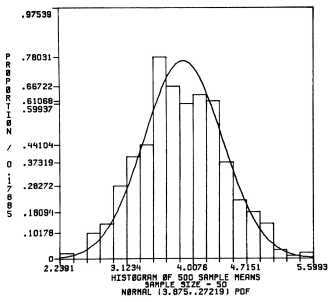

Figure 4: Gentleman (1977), p. 171

- **Empirical Cumulative Distribution Function Plots**
  - A plot used to estimate the cdf from given data points.

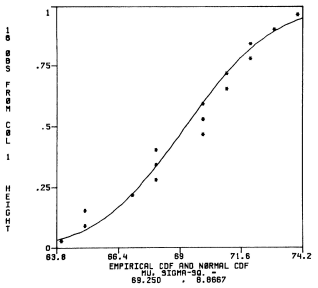

Figure 5: Gentleman (1977), p. 171

- **Probability Plots**
  - Similar to Empirical Cumulative Distribution Function Plots.
- **Regression**
  - "*The subject of regression lends itself nicely to teaching with plots*" - (Gentleman (1977), p. 172)
  - R Example: Anscombe data Gentleman (1977), Snell & Peterson (1992)

- **Likelihood Functions**
  - Assists students in making inference about the outcomes from increasing and decreasing the sample size.

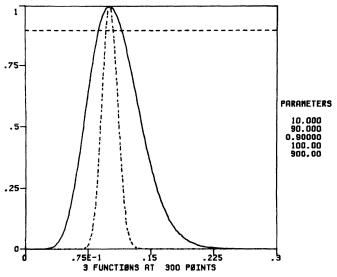

Figure 6: Gentleman (1977), p. 173

- **Enhanced Scatter Plots**
  - Easy for on the spot transposing of the data to show the students differences in regression when the axes are changed.

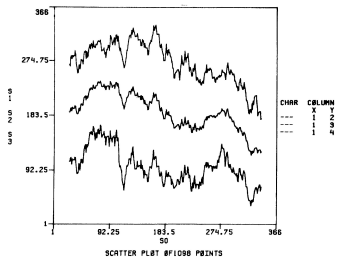

Figure 10.2

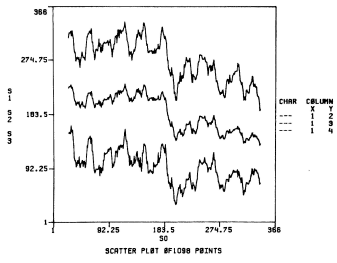

Figure 7: Gentleman (1977), p. 174

Graphics in the Statistics Classroom Today

- Courses I have taken:
  - **Stat 5200** - Probability plots are used to determine if the data is normal or needs a transformation.
  - **Stat 5720 and Stat 6550** - Plotting of maximum likelihood functions with various sample sizes.
  - **Stat 6100** - Histograms and goodness-of-fit to check residuals for linear regression (qqplots in R). Snell & Peterson (1992)
- Stat 1040 class I teach: (Note: the 1040 text contains few graphics and the addition of graphics in the course enhances student learning and understanding of the content)
  - **Histograms and Goodness-of-Fit**: When you get a histogram from some data what distribution does it appear to follow? If you know the distribution what can you do with that information?
  - **Regression**: What should the data look like in a scatter plot if you want to use linear regression?
  - **Confidence Intervals**: If I can find the confidence interval, how much of the data (usually on the normal curve) will be included? Gordon & Gordon (1992)
  - **Applets**: Applets are interactive simulations that the students operate to learn statistical concepts. With applets students (and teachers) do not have to be computer programming savvy to understand the concepts demonstrated.
    - Examples:
      - `http://www.math.usu.edu/~schneit/CTIS/`
      - `http://www.stat.tamu.edu/~west/applets/ci.html`
- **What if there is no technology available in the classroom?**
  - Histograms - Line students up to make a histogram of student heights.
  - Probability - Use a deck of cards to demonstrate various concepts with probability (i.e. with or without replacement, counting principles, etc.) and create charts to list possible outcomes.

- Law of averages - Divide the students into groups to flip coins and record the number of head and tails and calculate the difference between the observed value and expected value.
- Probability Histograms and Central Limit Theorem - Distribute bags with numbers (or dice) and have the groups of students draw numbers from a bag tally the sum and the product. After about 25 - 30 draws the chart with the sum will have a shape of a normal distribution.

- **Hypotheses of why Gentleman's suggestions may or may not be as prevalent in the classrooms of today:**

  (i) Gentleman's suggestions have been improved upon. Current movement in statistics education is to have more hands on learning with technology and group work.

  (ii) The statistical computer programs Gentleman suggested are a thing of the past. Most students (and some teachers) are not programming savvy, new programs have been implemented that are more user friendly and teach the concepts.

  (iii) Various technological tools are more prevalent today than in the 1970's (i.e. computers, calculators, etc.) and subsquently graphics are all over the place (for good and bad).

Links to R-code:

- `http://www.math.usu.edu/~symanzik/teaching/2009_stat6560/RDataAndScripts/allred_brittany_project1_rcode.R`

References

Gentleman, J. F. (1977), 'It's All a Plot (Using Interactive Computer Graphics in Teaching Statistics)', *The American Statistician* **31**(4), 166–175.

Gordon, F. S. & Gordon, S. P. (1992), Sampling + Simulation = Statistical Understanding: Computer Graphics Simulations of Sampling Distributions, *in* F. Gordon & S. Gordon, eds, 'Statistics for the Twenty–First Century (MAA Notes, Number 26)', The Mathematical Association of America, Washington, D.C., pp. 207–216.

Snell, J. L. & Peterson, W. P. (1992), Does the Computer Help us Understand Statistics?, *in* F. Gordon & S. Gordon, eds, 'Statistics for the Twenty–First Century (MAA Notes, Number 26)', The Mathematical Association of America, Washington, D.C., pp. 167–188.

Tijms, H. (1992), Exploring Probability and Statistics Using Computer Graphics, *in* F. Gordon & S. Gordon, eds, 'Statistics for the Twenty–First Century (MAA Notes, Number 26)', The Mathematical Association of America, Washington, D.C., pp. 189–197.