Visual Inference for Regression Parameters

Mahbubul Majumder, Heike Hofmann and Dianne Cook

Department of Statistics, Iowa State University, Ames, IA

Iowa State University

Working Group
Statistical Graphics

Introduction

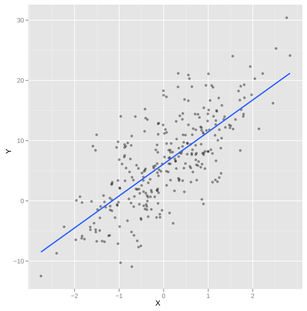
- Statistical graphics have been used for
  - exploratory data analysis.
  - model checking and diagnostics.
- Can we use them for Inference?
- Buja et al (2009)
  - Introduced method to test the significance of findings.
  - Their focus was on testing overall model fitting.
- Some times we are particularly interested in model parameters.
- This presentation focuses on testing parameters related to regression models.

Visual Statistical Inference

Visual Test Statistic

Unlike classical hypothesis testing the statistic in visual inference is not a single value, but a plot that is appropriately chosen to describe the parameter of interest.

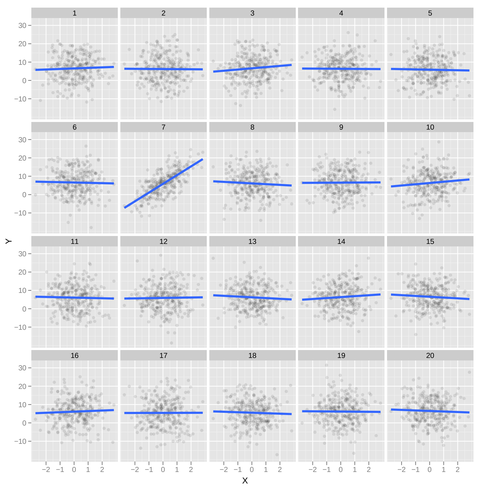
The visual test statistic shown here indicates the existence of positive slope.



Lineup Plot

A lineup plot is a layout of m visual statistics, consisting of

- $m - 1$ plots simulated from the model specified by H_0 (null plots) and
- the test statistic produced by plotting the observed data, possibly arising from H_1.

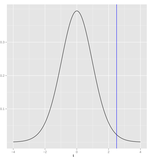
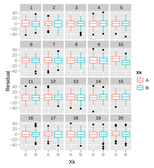


Can you identify the observed plot in this lineup?

We reject H_0 if the observed plot is identified.

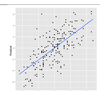
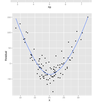
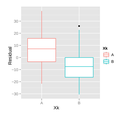
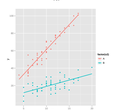
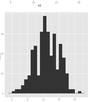
More than one person can evaluate a lineup.

Comparison with traditional inference

Steps	Mathematical Inference	Visual Inference
Hypothesis	$H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$	$H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$
Test statistic	$T(y) = \frac{\hat{\beta}}{se(\hat{\beta})}$	$T(y) =$ 
Null Distribution	$f_{T(y)}(t)$; 	$f_{T(y)}(t)$; 
Reject H_0 if	observed T is extreme	observed plot is identifiable

Visual Test Statistics Related to Regression Model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + ... + \epsilon_i \; ; \; \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

Null Hypothesis	Test Statistic	Description
$H_0 : \beta_k = 0$		Residual vs X_k Plot
$H_0 : X$ Linear		Residual vs X_k Plot
$H_0 : \beta_k = 0$; X_k categorical		Boxplot
$H_0 : \beta_k = 0$ (interaction with categorical X_k)		Scatter plot
H_0 : Model Fits		Histogram

Simulation Study

A simulation study was done for a simple linear model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta X_i + \epsilon_i$$

where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, X is categorical. $H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$

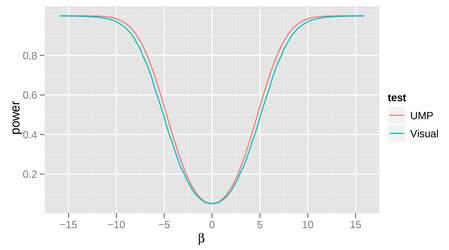
Combination of Parameters

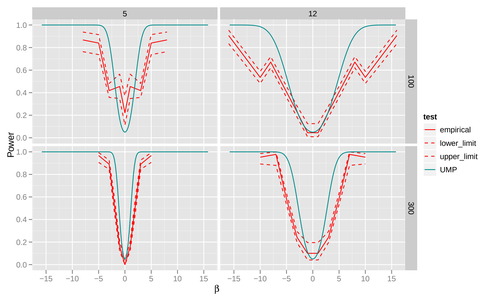
Sample size (n)	σ	values for β
100	5	0, 1, 3, 5, 8
	12	1, 3, 8, 10, 16
300	5	0, 1, 2, 3, 5
	12	1, 3, 5, 7, 10

- Three lineup plots were generated for each combination of parameters
- Amazon Mechanical Turk was used to recruit people
- Each of the 324 People evaluated 10 different lineups

Power of the Visual Test

Expected Power of visual test and the power of UMP test for sample size n = 100 and σ = 12.



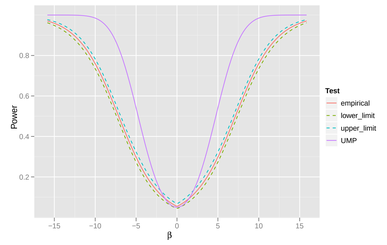


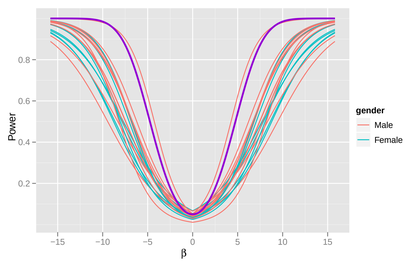
Observed Power of visual test, pointwise 95% confidence limits and the power of UMP test for sample size n = 100, 300 and σ = 12, 5.

Power Estimated from Model

A generalized mixed linear model fitted to Amazon Mechanical turk experiment data gives overall power as well as subject specific power.

Estimated power from model along with 95% confidence interval for sample size n = 100 and σ = 12. The corresponding power for UMP test is shown for comparison.





Estimated subject specific power from model for sample size n = 100 and σ = 12. The power curve for one subject is above (UMP) test power curve.

Ongoing and Future Work

- What if not normal?
  - Extend this study for generalized linear model.
- Apply the procedure with real data.
- Conduct survey
  - Examine the other test statistics.
  - Asses the sensitivity of power to modeling conditions.
  - Discover the most effective specification of a plot.

Acknowledgements: This work is funded by NSF grant DMS 1007697