Four ways to write up a model.

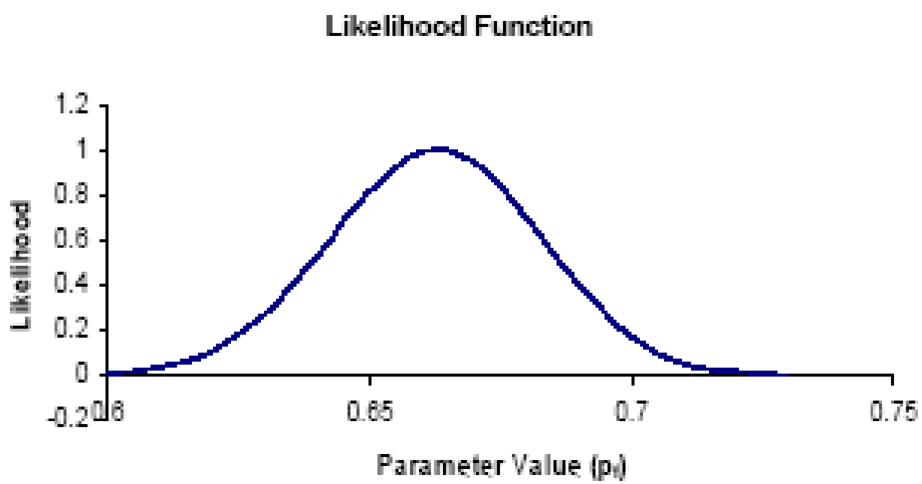
- English – Likelihood is the probability of observing the data given a parameter.

- Analysis (Equations)

$$\mathcal{L}(\theta \mid x_1, x_2, x_3 \ldots x_n) = \prod_{i=1}^{n} cg(x_i \mid \theta)$$

$$\mathcal{L}(p_f \mid data) = \prod_{i=1}^{4} p_f^{\,x_i}(1 - p_f)^{n_i - x_i}$$

–.



- Simulations
- theta<-seq(40)/40
- plot(theta,dbinom(5,10,theta))

The 3 essentials for statistical models

- A predictive model for the mean response as a function of explanatory variables.
- A description of the variation around the mean, i.e. a distributional assumption.
- The variation among studies of a similar form, i.e. the use of meta-analytic or hierarchical random effects or Bayesian methods.

3 Components of a GLIM

- A random component: this describes the variation in the data not explained by the functional model.
- A systematic component: a *linear* predictor from the explanatory variables.
- A link function that expresses the linear prediction in the scale of the data, i.e. the random component.

The Random Component or Error Distribution

- You must decide what type of variation is observed in your data. The four basic examples are:
- A. Normal
- B. Gamma
- C. Poisson
- D. Binomial

Common Error Distributions in Generalized Linear Models

Distribution	Probability Density	Variance Function

Normal

canonical link=indenity

range=real line

Gamma

canonical link=inverse

canonical model – rate of process

range=positive real line

Poisson

canonical link=log

canonical model – log of mean

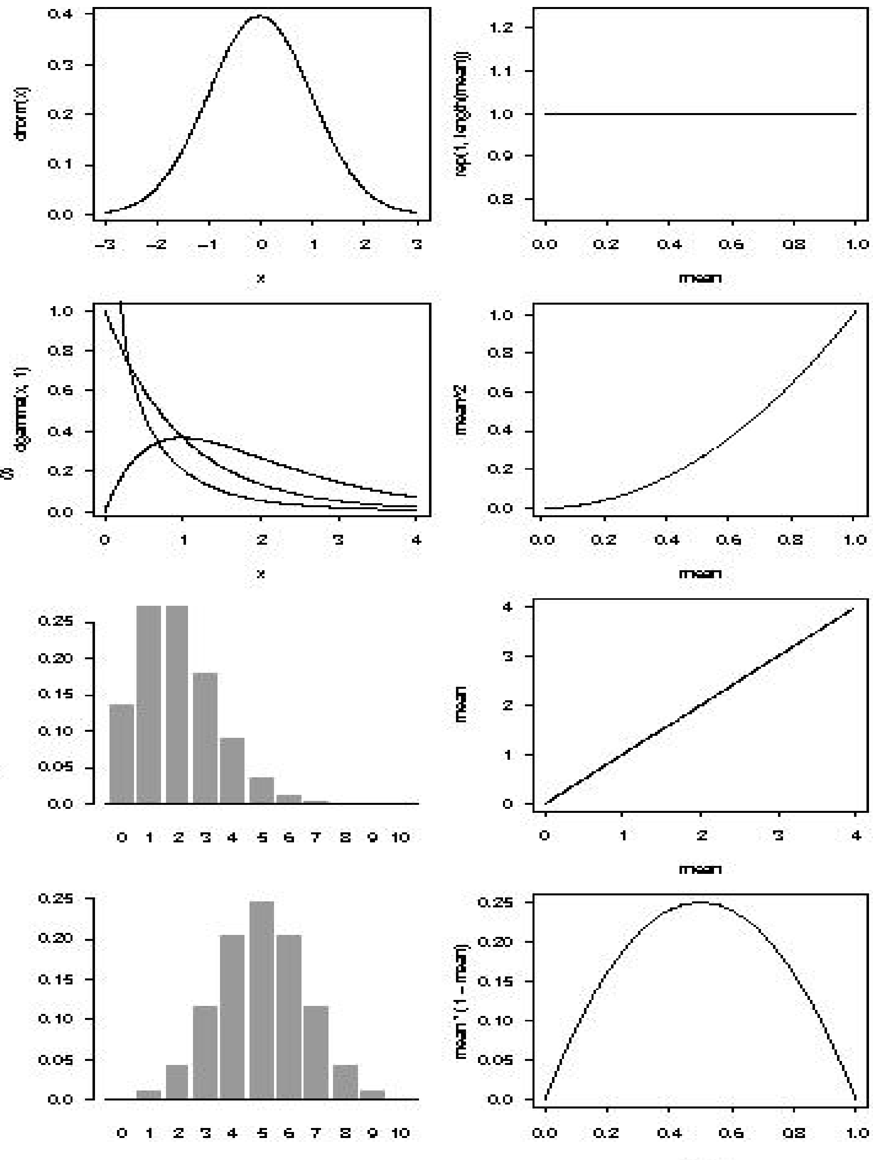
range=0,1,2,...

Binomial

canonical link=Logit

canonical model – log odds

range=0,1,3,...,N



psLABEL /autofs/auto/tloram/myers/teach/s postscript/graph.glm.dist.ps Mon Feb 14 23:32:42 AST 2005 graph.glm.dist()

Other possibilities

- Negative binomial (used in almost all cases where there is overdispersed count data)
- Inverse gamma (seldom used).

Systematic component

- This is the basic linear model that includes all of multiple linear regression and ANOVA's.
- You understand this.

Links Function

- GLIM's can be used to model a wide range of functional responses, much wider than can be used with standard linear models
- Each distribution has a canonical link, which has nice statistical properties.
- Other links can be used, the log link is often used with gamma model, and a variety of other links, e.g. the probit links
- can be used with the binomial distribution.
- As with linear models, a squared term, i.e. a quadratic, still keeps a link with the GLIM framework. This allows a wider range of models to be fit.

Common link functions used in Generalized Linear Models