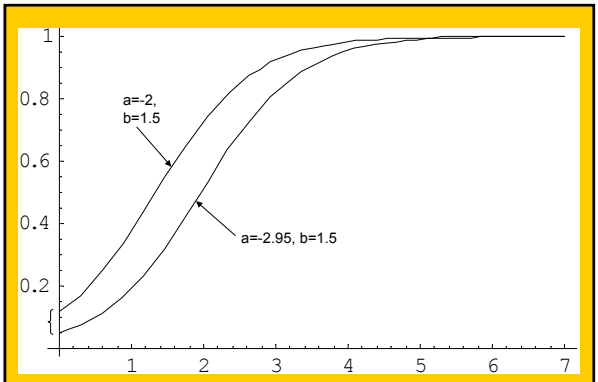
Logistic Regression

Modeling the Probability of an Event as a Function of an Independent Variable

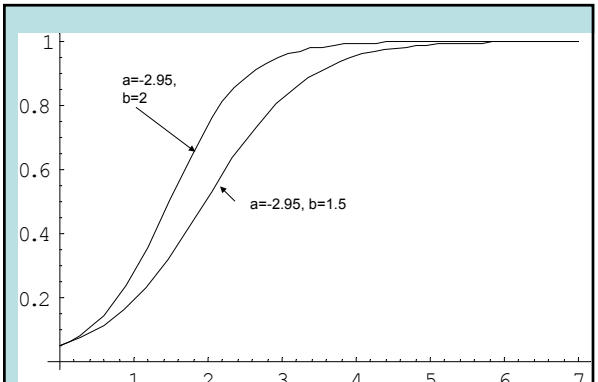
Equivalent Models

$$\pi = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$$

$$\ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta x$$



Logistic with Different alpha's



Logistic with different Betas

Logistic Regression

- Dichotomous response variable
- Predictors can be either qualitative or quantitative
- We can have many predictors or just one

Multiple Logistic Regression

$$\log it(\pi) = \alpha + \beta_1 x_1 + \ldots + \beta_p x_p$$

Alternative Models

$$\pi = \frac{e^{\alpha+\beta_1 x_1+\ldots+\beta_p x_p}}{1+e^{\alpha+\beta_1 x_1+\ldots+\beta_p x_p}}$$

$$\ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 x_1 + \ldots + \beta_p x_p$$

Survey Data

In favor of affirmative action (yes, no)
Political ideology(1 = very liberal, 2 =liberal,
3 = slightly liberal, 4 = moderate,
5 = slightly conservative, 6 = conservative,
7 = very conservative)
Gender (female, male)

Table of ge by pi

ge	pi						
Frequency Percent Row Pct Col Pct	1	2	3	4	5	6	Total
f	4 9.09 14.81 57.14	13 29.55 48.15 65.00	3 6.82 11.11 60.00	6 13.64 22.22 75.00	0 0.00 0.00 0.00	1 2.27 3.70 100.00	27 61.36
m	3 6.82 17.65 42.86	7 15.91 41.18 35.00	2 4.55 11.76 40.00	2 4.55 11.76 25.00	3 6.82 17.65 100.00	0 0.00 0.00 0.00	17 38.64
Total	7 15.91	20 45.45	5 11.36	8 18.18	3 6.82	1 2.27	44 100.00

Model Information	
Data Set	WORK.SURVEY
Distribution	Binomial
Link Function	Logit
Response Variable (Events)	aaa
Response Variable (Trials)	n
Observations Used	60
Number Of Events	44
Number Of Trials	60

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	57	54.2111	0.9511
Scaled Deviance	57	54.2111	0.9511
Pearson Chi-Square	57	56.3794	0.9891
Scaled Pearson X2	57	56.3794	0.9891
Log Likelihood		-27.1056	

Analysis Of Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		1	2.3258	0.8384	0.6825	3.9691	7.69	0.0055
ge	f	1	1.4351	0.7060	0.051	2.818	4.13	0.0421
ge	m	0	0.0000	0.0000	0.000	0.000	.	.
pi		1	-0.5822	0.2113	-0.996	-0.168	7.59	0.0059
Scale		0	1.0000	0.0000	1.000	1.000		

SAS Setup

- **data** survey;
- input subj ge $ ag hi co dh dr tv sp ne ah ve $ pa $ pi
- re ab $ aa $ ld $;
- if aa='y' then aaa=**1**; if aa='n' then aaa=**0**;
- n=**1**; cards;
- 1 m 32 2.2 3.5 0 5.0 3 5 0 0 n r 6 2 n n y
- 60 f 22 3.4 3.0 650 4 8 16 7 1 n i 4 1 y y y
- ;
- **proc genmod** data=survey; class ge;
- model aaa/n = ge pi /dist=bin link=logit;**run**;

AZT and Aids Symptoms

```
data aids;
 input race $ azt $ y n @@;
cards;
white yes 14 107 white no 32 113 black yes 11 63 black no 12 55
;
proc genmod; class race azt;
model y/n= race azt / dist=bin link=logit residuals; run;
```