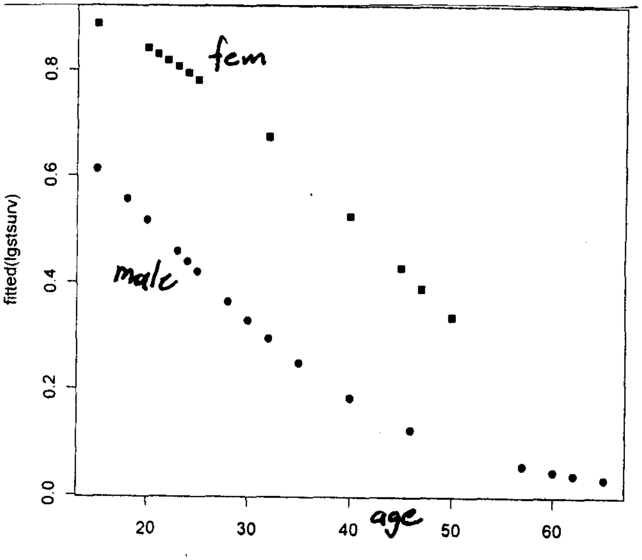
Logistic Regression Two Predictors, one measured, one dichotomous

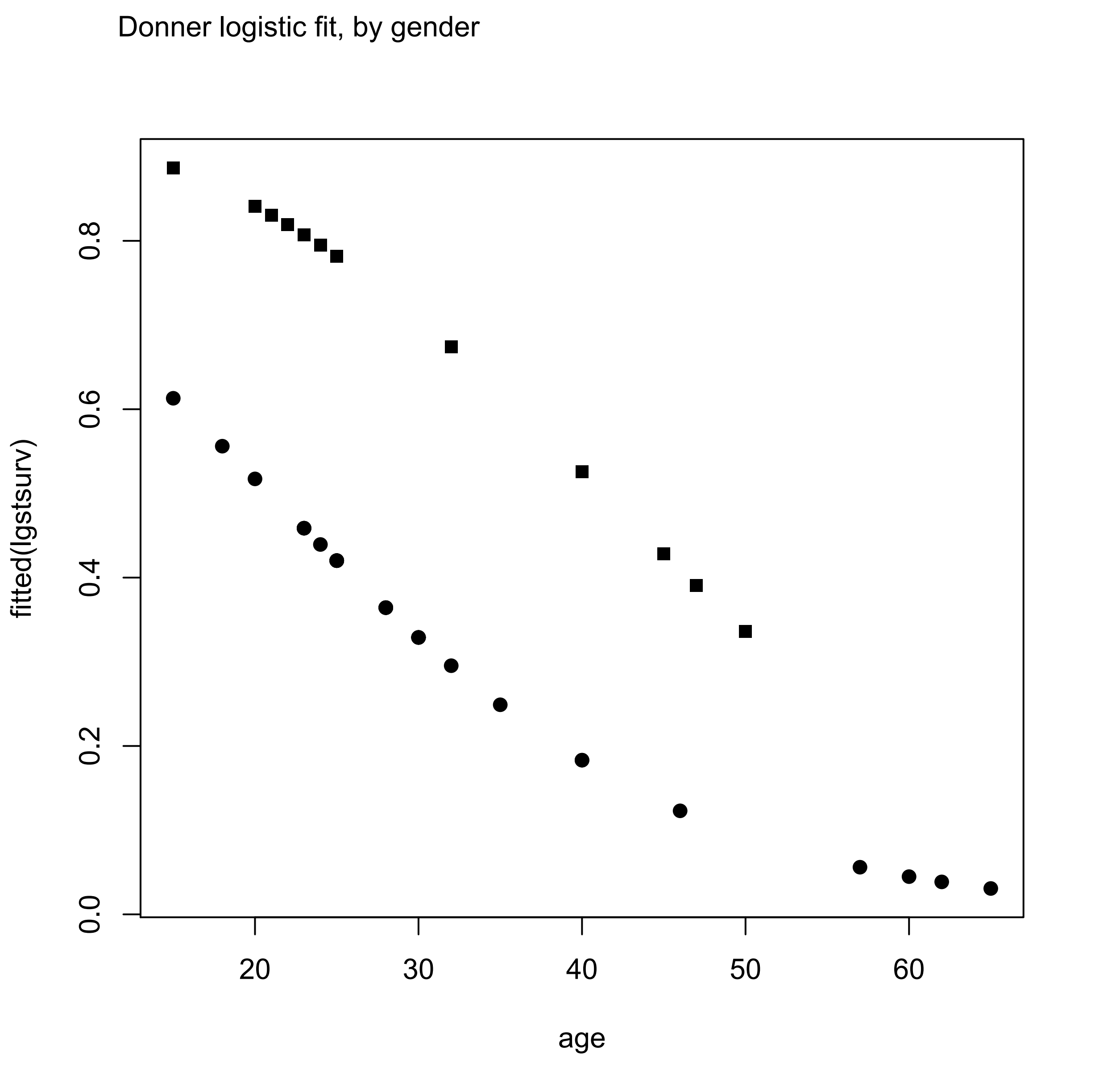
```
Survival in the Donner Party --Logistic Regression
File donner.dat contains the ages and sexes of the adult (over 15 years) survivors
and non-survivors of the party. These data were used by an anthropologist to study
the theory that females are better able to withstand harsh conditions than are males
(Data from D. K. Grayson, 1990, "Donner Party Deaths: A Demographic Assessment,"
Journal of Anthropological Research 46: (1990): 223—42.)
> donner = read.table(file="D:\\stat141\\donner.dat", header = T) #cols age male survival
> table(survival, male) |> prop.table(table(survival, male),2)
        male            |           male                > # females 2/3 survive, males 1/3 surv
survival  0  1          |survival          0         1  > #relrisk being male = 2
       0  5 20          |       0 0.3333333 0.6666667
       1 10 10          |       1 0.6666667 0.3333333
> smor = oddsratio(donsm, log = F)  > summary(smor)  Odds Ratio [1,]      0.25
> confint(smor)                    > confint(smor, level = .99)
          lwr       upr                      lwr      upr
[1,] 0.07007981 0.8918403          [1,] 0.04699297 1.329986
> lgstsurv = glm(survival ~ age + male, family = binomial)
> summary(lgstsurv)
Call: glm(formula = survival ~ age + male, family = binomial)
Coefficients:               Estimate Std. Error z value   Pr(>|z|)
             (Intercept)  3.23041    1.38686   2.329    0.0198 *
             age         -0.07820    0.03728  -2.097    0.0359 *
             male        -1.59729    0.75547  -2.114    0.0345 *
> cbind(age, male,predict.glm(lgstsurv), fitted(lgstsurv))
   age male  log-odds     prob surv  > #predict gives log-odds, fitted gives prob success
1   23    1 -0.16557336 0.45870097          type = "response",
2   40    0  0.10225091 0.52554048      > plot(age, fitted(lgstsurv),pch = male + 15)
3   40    1 -1.49504259 0.18316607
4   30    1 -0.71300187 0.32893588
5   28    1 -0.55659372 0.36433597
6   40    1 -1.49504259 0.18316607
7   45    0 -0.28876945 0.42830515
8   62    1 -3.21553219 0.03858538
9   65    1 -3.45014441 0.03076455
10  45    0 -0.28876945 0.42830515
...
45
```

donsm = table(survival, male); refer to logit table 11/17; log-odds = 3.23 − .078A − 1.597M; $e^{-.078} = .925$ multiplicative change in odds for 1yr increase in age; fitted prob survival; odds = P/(1−P); p = odds/(odds+1)




```
> exp(-1.5973)
[1] 0.2024424 #odds 1/5 as large for male, constant age
> predict.glm(lgstsurv, type="response",
newdata = data.frame(male=0, age=20))   [1] 0.8410862
> predict.glm(lgstsurv, type="response",
newdata = data.frame(male=0, age=50))   [1] 0.3363082
```

$e^{\hat\beta}$ change from male = 0 to male = 1 in odds; odds 5; F50 .51; F20 5.3


Donner logistic fit, by gender
fitted(lgstsurv)
0.0
0.2
0.4
0.6
0.8
20
30
40
50
60
age

Donner ex

logistic fit

log-odds(survival) = 3.23 - .0782 age - 1.597 male

for males

log-odds(survival)

= 1.633 - .0782 age

for females

log-odds(survival)

3.23 - .0782 age

age response same, male/female diff overall.

odds(survival)

$= e^{(1.633 - .0782 age)}$

odds(survival)

$e^{(3.23 - .0782 age)}$

$prob = \frac{odds}{odds+1}$ gets logistic function as in 54

ancova for logistic model

Stat 141
11/30/06

logistic regression page: dichotomous outcome

Donner Party revisited

```
> donner = read.table(file="D:\\stat141\\donner.dat", header = T) #cols age male survival
> lgstsurv = glm(survival ~ age + male, family = binomial) # logistic regr fit,  12/1
> summary(lgstsurv)
Call: glm(formula = survival ~ age + male, family = binomial)
Coefficients:              Estimate Std. Error z value   Pr(>|z|)
             (Intercept)  3.23041    1.38686   2.329   0.0198 *
             age         -0.07820    0.03728  -2.097   0.0359 *
             male        -1.59729    0.75547  -2.114   0.0345 *
#glm says: log-odds survival = 3.23 - 0.078*age -1.597*male
 Using logistic fit  # odds = prob(1 - prob);  prob = odds/(odds + 1),see class 11/17
> exp(-1.5973) [1] 0.2024424      fem -> male
> # odds 1/5 as large for male survival compared to odds of male survival at any age
> exp(-.0782) [1] 0.9247795      age -> age+1
> # odds .925 as large for survival of 1-yr older for m or fem survival at any age
> exp(30*(-.0782)) [1] 0.0957514   age -> age+30
> # odds .096 as large for survival of 30-years older for m or fem survival, any age
compare fits--fitted probability of survival--for 20yr old 50yr old males and females
> #predict gives log-odds, fitted gives prob success (or predict w/ type="response") prob
> predict.glm(lgstsurv, type="response", newdata = data.frame(male=0, age=20))
[1] 0.8410862                                                     females
> predict.glm(lgstsurv, type="response", newdata = data.frame(male=0, age=50))
[1] 0.3363082
> predict.glm(lgstsurv, type="response", newdata = data.frame(male=1, age=20))
[1] 0.5172529                                                      males
> predict.glm(lgstsurv, type="response", newdata = data.frame(male=1, age=50))
[1] 0.09303878
> .841/.159;   .517/.483;   .336/.664;    .093/.907   #odds for fitted probabilities
[1] 5.2893    [1] 1.0704   [1] 0.5060    [1] 0.1025
> .506/5.29;     .103/1.07        #odds ratios for being 30 yr older (as above)
[1] 0.09565217  [1] 0.09626168                                         matches
 # odds ratio for male vs female survival at age 20 or age 50 = .2   matches
To make a chart of fits for the 45 observed values of age,male
> cbind(age, male,predict.glm(lgstsurv), fitted(lgstsurv)) see 12/1 output
aside:  if you wanted to fit logistic serately for males for ex
> lgstsurvm = glm(survival[male==1] ~ age[male==1], family = binomial)   subsetting
```

ex 1

	age	male	survival
01	23	1	0
02	40	0	1
03	40	1	1
04	30	1	0
05	28	1	0
06	40	1	0
07	45	0	0
08	62	1	0
09	65	1	0
10	45	0	0
11	25	0	0
12	28	1	1
13	28	1	0
14	23	1	0
15	22	0	1
16	23	0	1
17	28	1	1
18	15	0	1
19	47	0	0
20	57	1	0
21	20	0	1
22	18	1	1
23	25	1	0
24	60	1	0
25	25	1	1
26	20	1	1
27	32	1	1
28	32	0	1
29	24	0	1
30	30	1	1
31	15	1	0
32	50	0	0
33	21	0	1
34	25	1	0
35	46	1	1
36	32	0	1
37	30	1	0
38	25	1	0
39	25	1	0
40	25	1	0
41	30	1	0
42	35	1	0
43	23	1	1
44	24	1	0
45	25	0	1