April 12, 2015
Erik Parner

PhD Course in Basic Biostatistics
Spring 2015
A short solution

Question 1

The output of the logistic regression with the age group as explanatory variable contains an estimate of the odds for falling among "under 70 years" of 0.33 (95% confidence interval, CI: 0.21-0.54), and odds ratio for the groups

70-74 years:	0.93 (0.50-1.74)
75-79 years :	1.14 (0.61-2.13)
80-84 years:	1.40 (0.73-2.70)
85-89 years:	3.10 (1.55-6.18)
90 years .. :	2.46 (1.18-5.13)

as compared to the reference age group "under 70 years". The odds rate cannot be interpreted as a relative risk as the risk of falling is not a rare event. The odds of falling, and therefore also the risk of falling, generally increase with the age of the woman.

Figure 1 shows that it is reasonable to assume that the log-odds depend linearly on the woman's age. Figure 2 shows that it is also reasonable to assume that the log-risk depends linearly on the woman's age.

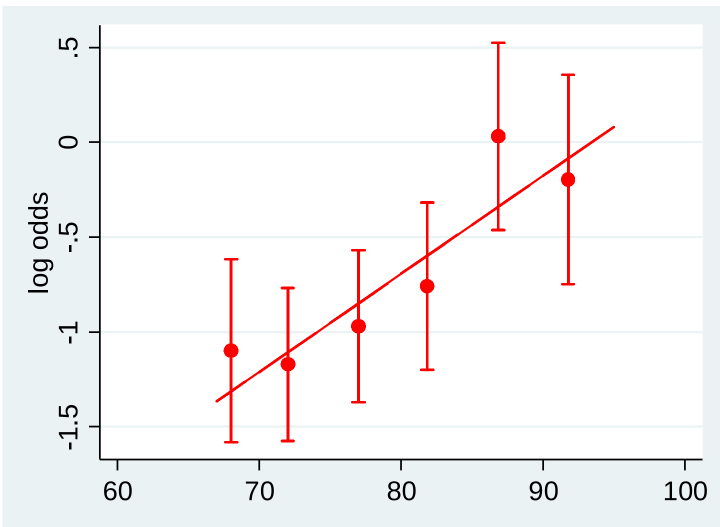

Figure 1 Log-odds against age group for women.

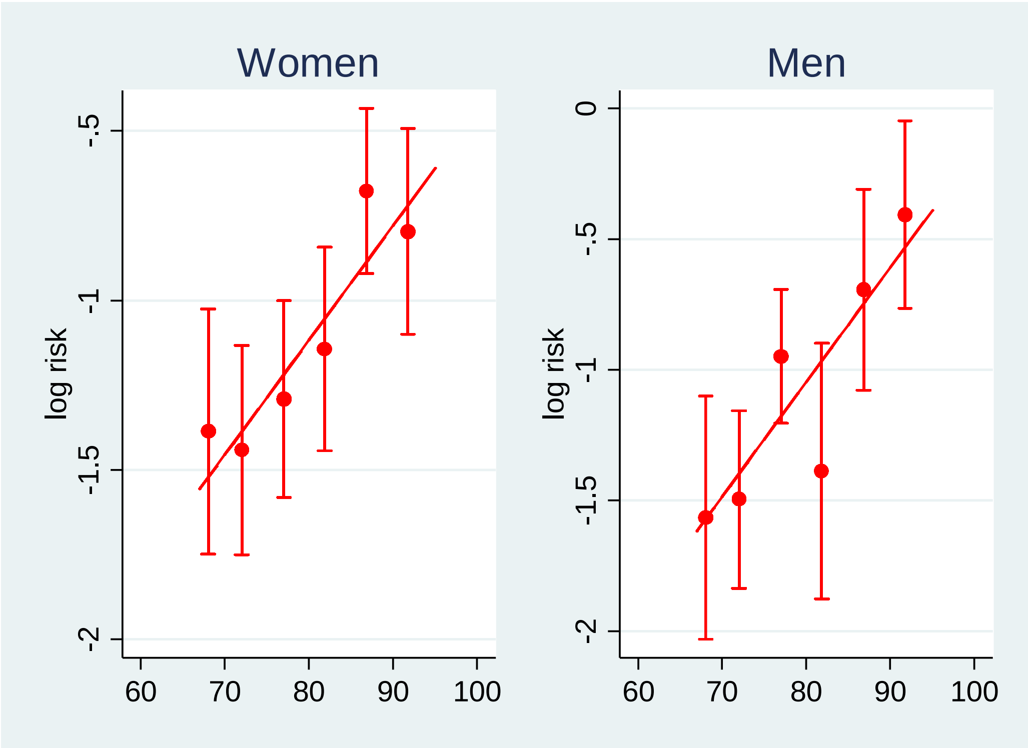

Figure 2 Log-risk against the age for women (left) and men (right).

Question 2

In the binary regression model for the relative risk the women of age 70 is chosen as reference persons. In this case the regression parameters in the model are

a. 0.23 (0.19-0.29) is the risk of falling for woman of age 70.

b. 1.03 (1.02-1.05) is the relative risk of falling for two women where one is 1 year older than the other.

c. 0.97 (0.71-1.32) is the relative risk of falling comparing men to women, both of age 70 years.

d. 1.01 (0.99-1.04) is the ratio of 1 years relative risk for men as compared to women.

The age dependency is not statistical significant different between men and women (p=0.42).

The age adjusted relative risk of falling comparing men to women is 1.07 (0.88-1.30).

Question 3

The risk of falling within each of the balance categories is

below 40: 0.71 (0.60-0.82)
40-41: 0.53 (0.40-0.65)
42-43: 0.47 (0.40-0.58)
44-45: 0.40 (0.28-0.52)
46-50: 0.33 (0.27-0.39)
51-60: 0.13 (0.10-0.17)

Chosen the “above 50” group as the reference we obtain the risk differences:

below 40: 0.57 (0.46-0.69)
40-41: 0.39 (0.26-0.52)
42-43: 0.34 (0.23-0.45)
44-45: 0.27 (0.14-0.39)
46-50: 0.19 (0.13-0.26)
51-60: 0.00 (reference).

Question 4

The Berg Balance index is plotted against the self-evaluated balance sum score in Figure 3. There is approximately a linear relationship between the two balance score. The Berg Balance index could be predicted from the estimates from the linear regression models with the intercept of 7.76 (0.45-15.07) and slope 0.73 (0.58-0.88). One could quantify the difference between the observed and predicted Berg Balance index – the residual – but we will not go into further analyses here.

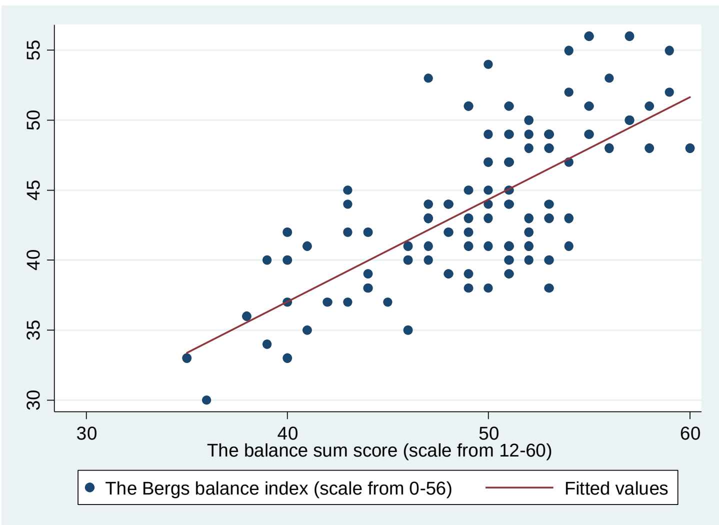

Figure 3 Scatter plot of the Berg Balance index against the self-evaluated balance sum score.

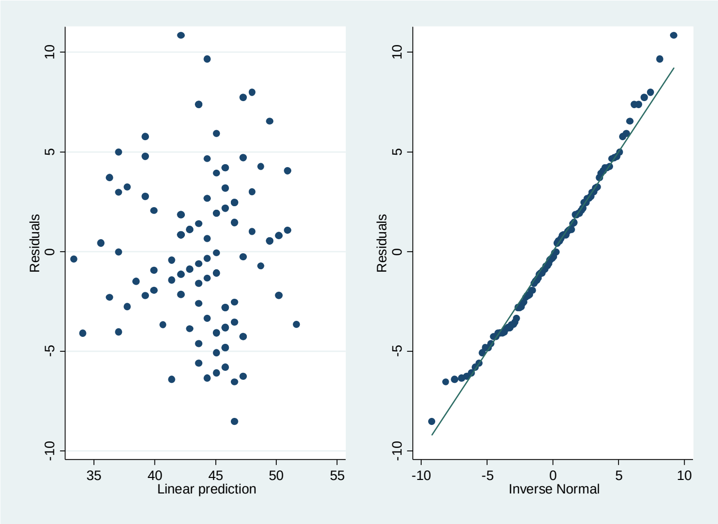

Figure 4 Plot of residual against the predicted values, and QQ plots of the residuals.

Question 5

The proportion of persons with a self-evaluated balance score below 48 among those with a low Berg Balance index (the sensitivity) is 0.57 (0.39-0.73), corresponding to 21 out of 37 persons.

The proportion of persons with a self-evaluated balance score 48 or higher among those with a high Berg Balance index (the specificity) is 0.83 (0.72-0.91), corresponding to 54 out of 65 persons.

Do-file

```
*******************************************************************
* Solution.do
* Solution for the exam Basic Biostatitics Spring.
* Erik Parner. 13-04-2015.
*******************************************************************

cd "D:\Teaching\BasicBiostat\Exam"

capture log close
log using "Solution.txt", replace text

use balance, clear

*******************************************************************
* Question 1.
*******************************************************************
recode age (min/69=1 "under 70 years") ///
        (70/74=2 "70-74 years")     ///
           (75/79=3 "75-79 years")     ///
           (80/84=4 "80-84 years")      ///
           (85/89=5 "85-89 years")     ///
           (90/max=6 "90 years or older"), generate(agegr)
* We shall only use agegr2 when plotting. There are several
* ways of constructing a variable with the mean age value
* within each age group, here is one short:
bysort agegr: egen agegr2=mean(age)

* Logistic model with agegr for womens.
logit fall i.agegr if(sex==0), or
predict logoddsgr, xb
predict se, stdp
gen lower=logoddsgr-1.96*se
gen upper=logoddsgr+1.96*se
logit fall c.age if(sex==0), or
predict logoddsline, xb
twoway (line logoddsline age if(sex==0),lco(red)) ///
      (rcap lower upper agegr2 if(sex==0) ,lco(red)) ///
         (scatter logoddsgr agegr2 if(sex==0), msy(O) mco(red))  ///
         , scale(1.5) legend(off) ytitle("log odds")
drop logoddsline logoddsgr se lower upper

* Binary regression for the relative risk with agegr for womens.
binreg fall i.agegr if(sex==0), rr
predict logoddsgr, xb
```

```
predict se, stdp
gen lower=logoddsgr-1.96*se
gen upper=logoddsgr+1.96*se
binreg fall c.age if(sex==0), rr
predict logoddsline, xb
twoway (line logoddsline age if(sex==0),lco(red)) ///
      (rcap lower upper agegr2 if(sex==0),lco(red)) ///
          (scatter logoddsgr agegr2 if(sex==0), msy(O) mco(red))  ///
          , scale(1.5) legend(off) ytitle("log risk") ///
          title("Women") name(graph1, replace)
drop logoddsline logoddsgr se lower upper

* Binary regression for the relative risk with agegr for men.
binreg fall i.agegr if(sex==1), rr
predict logoddsgr, xb
predict se, stdp
gen lower=logoddsgr-1.96*se
gen upper=logoddsgr+1.96*se
binreg fall c.age if(sex==1), rr
predict logoddsline, xb
twoway (line logoddsline age if(sex==1),lco(red)) ///
      (rcap lower upper agegr2 if(sex==1),lco(red)) ///
          (scatter logoddsgr agegr2 if(sex==1), msy(O) mco(red))  ///
          , scale(1.5) legend(off) ytitle("log risk") ///
          title("Men") name(graph2, replace)
drop logoddsline logoddsgr se lower upper
graph combine graph1 graph2
graph drop graph1 graph2


*******************************************************************
* Question 2.
*******************************************************************
generate age70=age-70
binreg fall i.sex##c.age70 , rr
binreg fall i.sex c.age70 , rr


*******************************************************************
* Question 3.
*******************************************************************
recode balance (min/39=1 "below 40") ///
         (40/41=2 "40-41")     ///
            (42/43=3 "42-43")     ///
            (44/45=4 "44-45")     ///
            (46/50=5 "46-50")      ///
```

```
            (51/60=6 "51-60"), generate(balancegr)
tabu balancegr fall, row
binreg fall ibn.balancegr , rd nocons
binreg fall ib6.balancegr , rd


*******************************************************************
* Question 4.
*******************************************************************
scatter berg balance
scatter berg balance, jitter(1)
regress berg balance
twoway (scatter berg balance) ///
      (lfit berg balance)
predict fit if e(sample), xb
predict res if e(sample), res
scatter res fit, name(graph1,replace)
qnorm res, name(graph2,replace)
graph combine graph1 graph2
graph drop graph1 graph2


*******************************************************************
* Question 5.
*******************************************************************
gen lowberg=(berg<42) if(berg<.)
gen highberg=(berg>=42) if(berg<.)
gen lowbalance=(balance<48) if(balance<.)
gen highbalance=(balance>=48) if(balance<.)
tabu lowbalance lowberg
ci lowbalance if(lowberg==1), binom
ci highbalance if(highberg==1), binom


log close
```