April 21, 2016
Erik Parner

PhD Course in Basic Biostatistics
Spring 2016

One (potential) solution

1. Compare the birth weights in the two groups without correction for gestational age. In particular, compute an estimate and a 95%-confidence interval for the difference in expected birth weights.

QQ-plot shows that the birth weight is approximately normal (Figure 1). The difference in birth weigh between the fish oil group and the control group is 102 g (95% CI: 9-194 g, p=0.03 using an unpaired t-test).

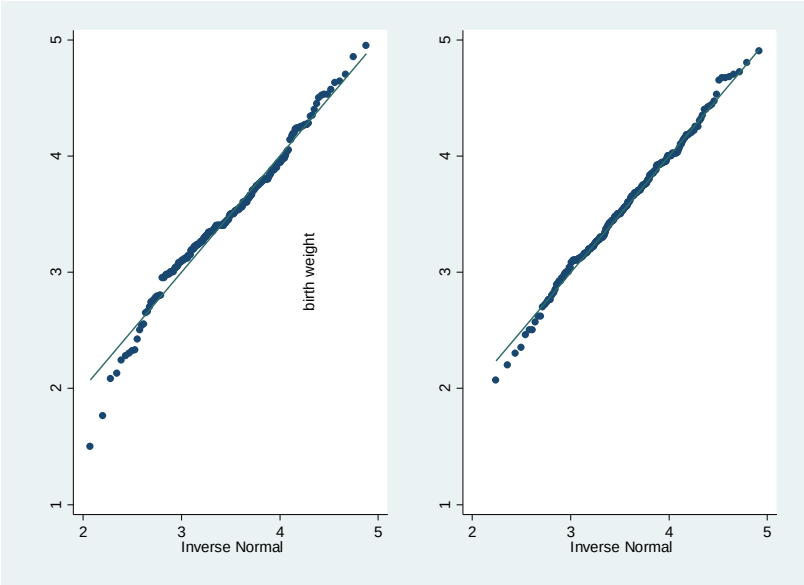

Figure 1 QQ-plot

2. Compare the gestational ages in the two groups.
The women tended to give birth at a later gestational age in the fish oil group compared to the control group (p=0.005 using a Chi-2 test, and p=0.0001 using Spearman test).

3.1. Establish a linear regression model of birth weight as dependent variable and gestational age as independent variable.
The relationship between of birth weight and gestational age seem approximately linear in both groups (Figure 2).

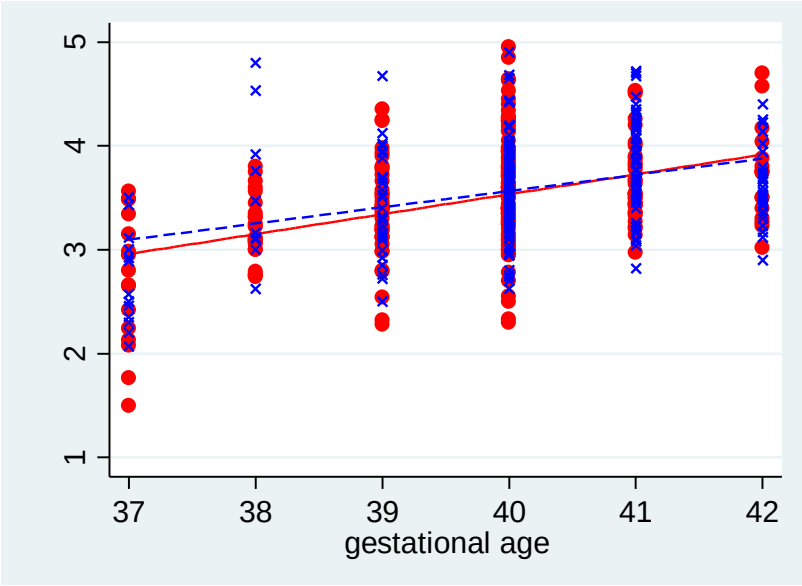

Figure 2 A linear regression of birth weight on gestational age in the control group (red) and the fish oil group (blue).

3.2. In the regression model, let the intercept correspond to gestational age 37 weeks, and compute estimates and 95%-confidence intervals for all model parameters. What is the interpretation of the regression model and each of its parameters?
For the control group the intercept is 2.96 kg (CI: 2.81-3.11 kg) and the slope is 0.19 (CI: 0.14-0.24 kg) and for the fish oil group the intercept is 3.10 (CI: 2.93-3.26 kg) and slope 0.16 (CI: 0.11-0.21). The intercept describes the mean birth weight at 37 weeks of gestation and the slope describes the average difference in birth weight between two women giving birth with a gestational age difference of 1 weeks.

3.3. Is the estimated slope significantly different from zero?
The slopes are significantly difference from zero ($p<0.001$ in both groups).

3.4. Based on the estimated model, calculate an interval that contains the birth weight of 95% of future newborns born at 40 weeks.
The 95% prediction intervals for the birth weight of children born in weeks 40 is 2.60-4.47 in the control group and 2.64-4.49 in the fish oil group.

4.1. Can the slopes in the two treatment groups be assumed equal?
The slope in the control group and fish oil group can be assumed equal ($p=0.32$).

4.2. Compare the birth weights in the two groups when correcting for gestational age.
The difference in birth weight between the fish oil group and the control group is 35g (CI:(-51g)-120g)), adjusting for gestational age.

5.1. Compare the risk of post term delivery in the two groups.
Since the risk of post term delivery is below 10% we interpret the odds ratio as a relative risk. The relative risk for post term delivery in the fish oil group compared to the control group is 1.76 (CI: 0.92-3.35).

5.2. Compare the risk of post term delivery in the two groups when correcting for birth weight in two groups; ≤3 kg, >3 kg.
Note there is only on post term child with a birth weight below 3 kg. Restricting on birth weights above 3 kg one obtain a relative risk for post term delivery delivery in the fish oil group compared to the control group is 1.63 (CI: 0.85-3.14). Assuming no effect measure modification between fish oil/ control group and birth weight group one obtain a relative risk of 1.71 (CI: 0.90-3.28).

6. Write a brief summary of your analyses and conclusions.
…

Do-file

```
*****************************
* Obligatory exercise.
*****************************

cd "D:\Teaching\BasicBiostat\Exam"

clear

capture log close
log using "Spring2016 solution", text replace

use faroe.dta, clear

des

* Question 1.
qnorm weight if(group==1), name(graph1,replace)
qnorm weight if(group==2), name(graph2,replace)
graph combine graph1 graph2, name(graph3)
graph export Figure1.wmf,replace
graph drop graph1 graph2 graph3
sdtest weight, by(group)
ttest weight, by(group)

* Question 2.
tabu ga group, chi2 col
tabu ga group, exp
spearman ga group

* Question 3.
twoway   ///
   (scatter  weight ga if group==1, msy(O)    mco(red) )  ///
   (lfit     weight ga if group==1, lpa(1)    lco(red) )  ///
   (scatter  weight ga if group==2, msy(X)    mco(blue) ) ///
   (lfit     weight ga if group==2, lpa(dash) lco(blue) ) ///
   , legend(off) ytitle("Birth weight (gram)") scale(1.5) name(graph1)
graph export Figure2.wmf,replace

* Control group.
generate ga37=ga-37
regress weight ga37 if group==1
predict fit if e(sample),xb
predict res if e(sample),res
scatter res fit, mcolor(red) yline(0) name(graph1,replace) scale(1.5)
scatter res ga37, mcolor(red) yline(0) name(graph2,replace) scale(1.5)
```

```
drop fit res
lincom _cons+3*ga37
disp r(estimate)-1.96*e(rmse) , r(estimate)+1.96*e(rmse)

* Fishoil group.
regress weight ga37 if group==2
predict fit if e(sample),xb
predict res if e(sample),res
scatter res fit, mcolor(red) yline(0) name(graph3,replace) scale(1.5)
scatter res ga37, mcolor(red) yline(0) name(graph4,replace) scale(1.5)
drop fit res
lincom _cons+3*ga37
disp r(estimate)-1.96*e(rmse) , r(estimate)+1.96*e(rmse)
graph combine graph1 graph2 graph3 graph4, ysize(3) name(graph5) col(2)
graph export Figure3.wmf,replace
graph drop graph1 graph2 graph3 graph4 graph5

* Question 4.
regress weight b1.group##c.ga37
regress weight b1.group c.ga37

* Question 5.
* The risk for postterm is less than 10%, so vi can use OR=RR.
tabu ga
gen group2=group-1
gen postterm=(ga>=42)
gen weightgr=(weight<=3)
tabu weightgr postterm
* Note: only one postterm is under 3 kg.
cs postterm group2 ,or
* The crude OR without adjustment for birth weight group.
tabu postterm group
logit postterm b1.group , or
bysort weightgr: tabu postterm group
* One can adjust my restriction on birth weight weight>3:
logit postterm b1.group if weightgr==0 , or
* Assuming no interaction between group and birth weight group:
logit postterm b1.group b0.weightgr , or

log close
```