lecture 4.3 part 3

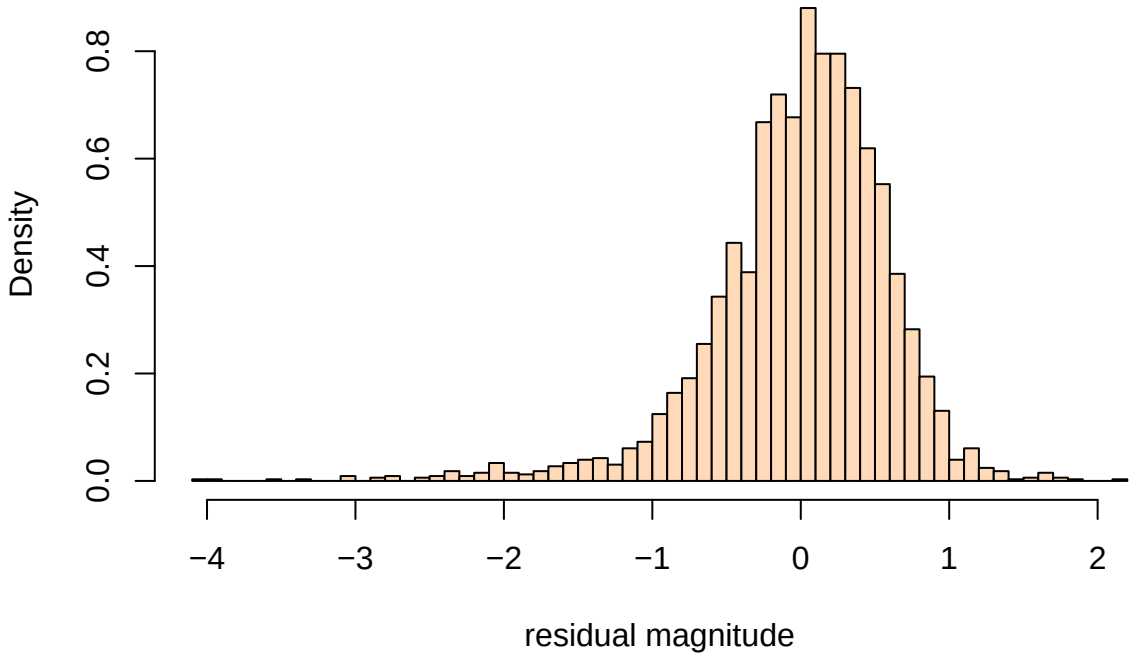
preliminaries

```
#clear workspace
rm(list=ls())

# load data
data = read.csv("~/Desktop/wages1.csv")

# build variables
wage = data$wage
school = data$school
exper = data$exper
```

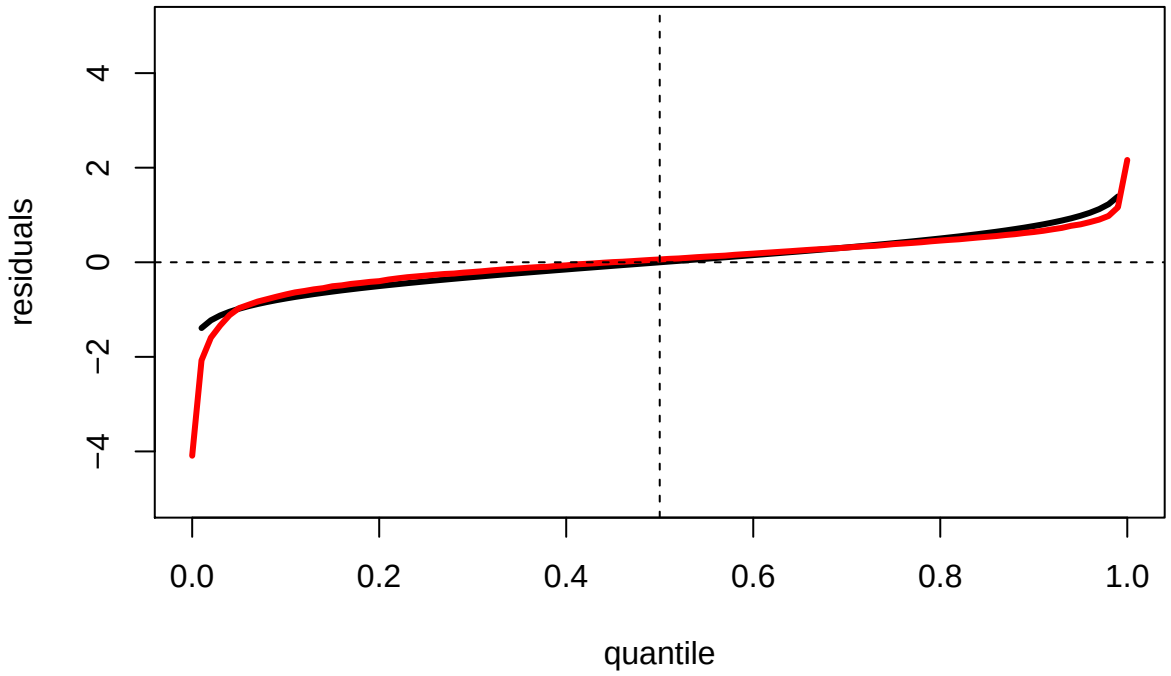
Fit linear model using lm() and look at residuals

NOTE: We will see in a later lecture why we can use regular OLS to estimate the beta-hats.

```
# compute the errors using linear regression using linear least squares (LLS)
# short-cut since we know that the MAP beta estimates are the ones given by the
# standard LLS
schoolC = school - mean(school)
model = lm(log(wage) ~ schoolC)
modelResid = resid(model)

#histogram of residuals
hist(modelResid,
     col="peachpuff",
     border="black",
     prob = TRUE, # show densities instead of frequencies
     xlab = "residual magnitude",
     main = "",
     breaks = 50)
```

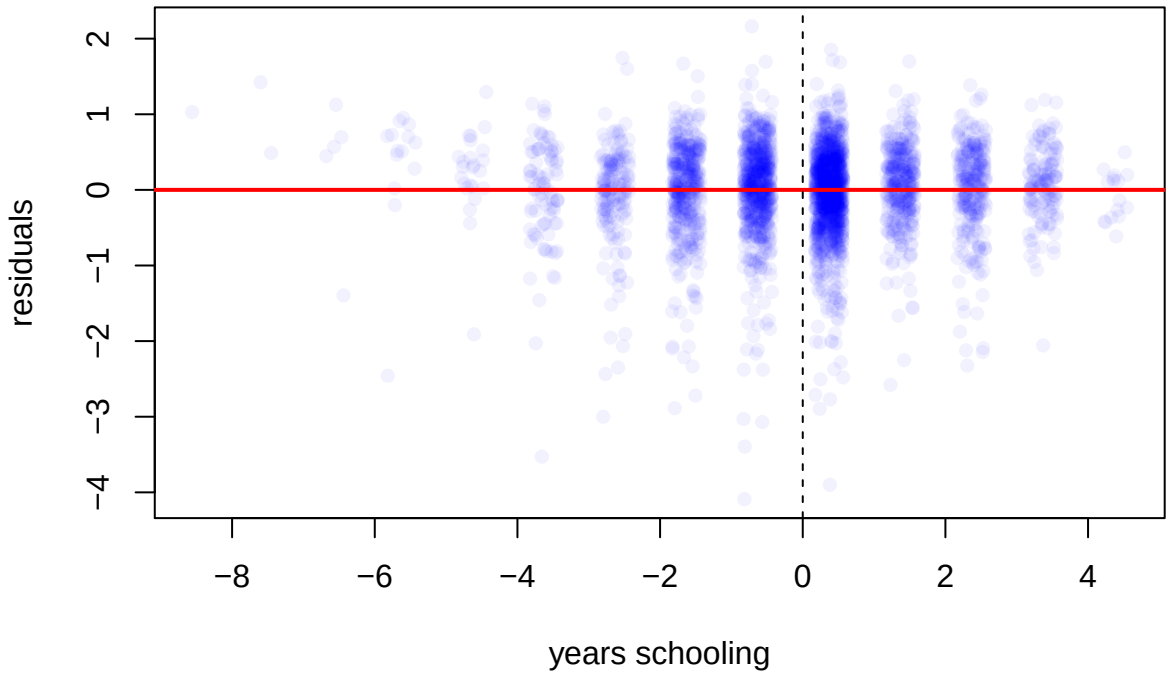



```
#normality of residuals
x = seq(0,1,by=0.01)
plot(x, qnorm(x, mean(modelResid), sd(modelResid)),
     type="l", ylim=c(-5,5), lwd = 3,
     xlab = "quantile", ylab = "residuals")
points(x, quantile(modelResid, probs=x), type="l", col="red", lwd=3)
abline(v=0.5, lty=2)
abline(h=0, lty=2)
```

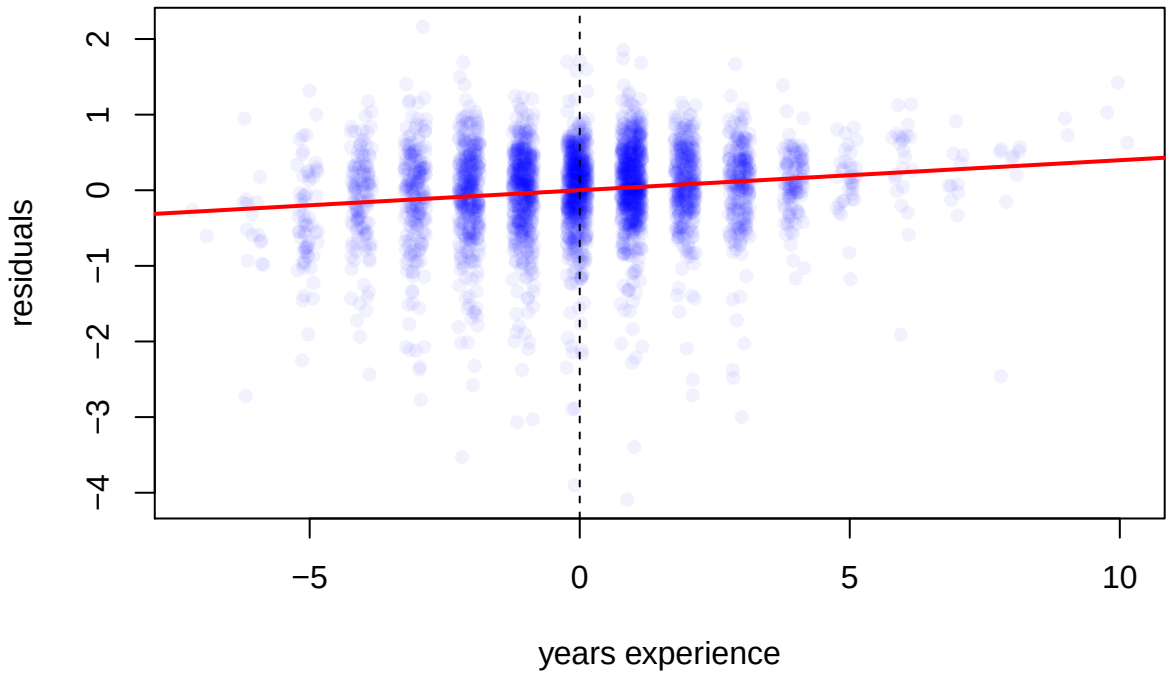



```
# plot residuals vs schooling
plot(jitter(schoolC,1), modelResid,
     col=rgb(red=0.0, green=0.0, blue=1.0, alpha=0.05),
```

```
     pch = 16,
     xlab = "years schooling", ylab = "residuals")
abline(lm(modelResid ~ schoolC), lwd = 2, col="red")
abline(v=0, lty=2)
```




```
# plot residuals vs experience
experC = exper - mean(exper)
plot(jitter(experC,1), modelResid,
     col=rgb(red=0.0, green=0.0, blue=1.0, alpha=0.05),
     pch = 16,
     xlab = "years experience", ylab = "residuals")
abline(lm(modelResid ~ experC), lwd = 2, col="red")
abline(v=0, lty=2)
```


residuals
2
1
0
−1
−2
−3
−4
−5
0
5
10
years experience