Homework 6

Aji John

Winter Quarter 2018

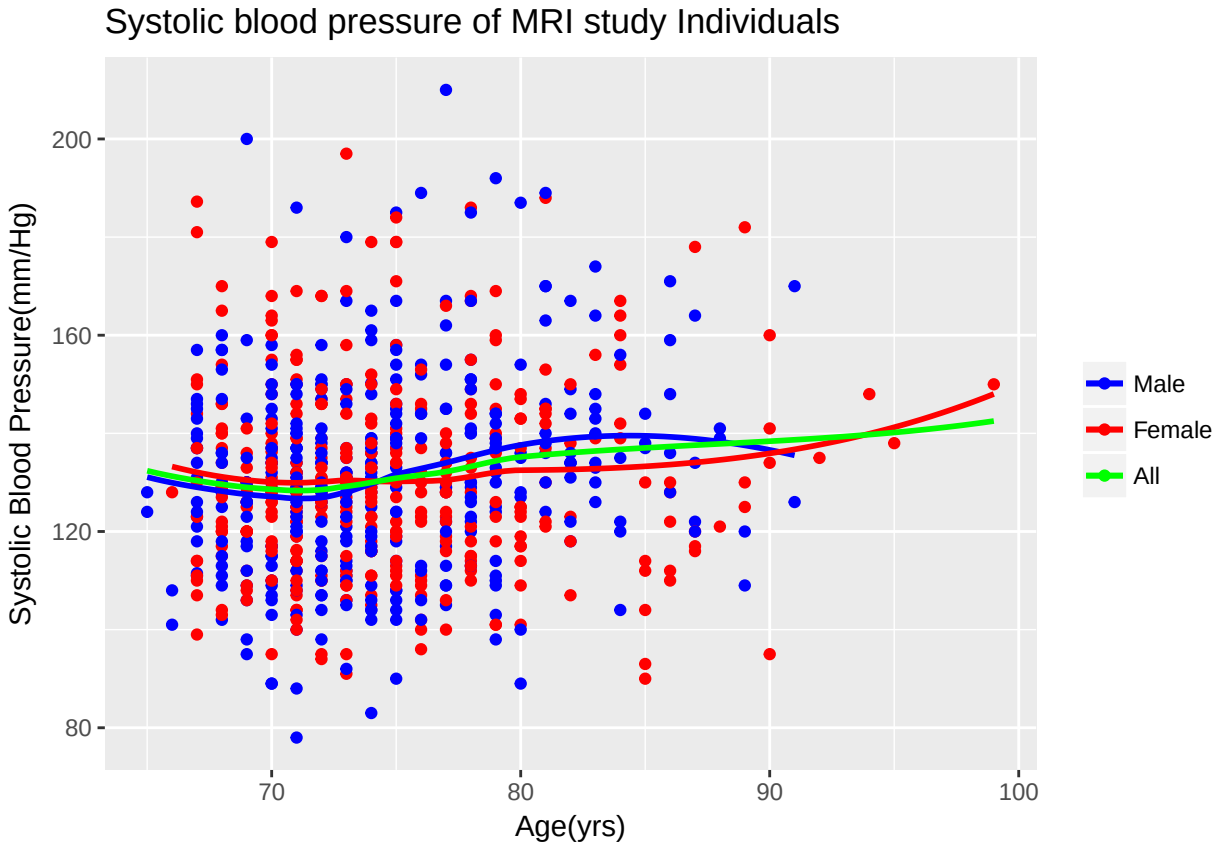
We are interested in assessing associations among systolic blood pressure, age, sex, and race.The sample is of generally healthy elderly subjects from four U.S. communities.

Question 1

We are interested in examining how mean systolic blood pressure varies by age and sex.

(a)

Create a scatterplot of systolic blood pressure versus age. Use different symbols and/or colors for each sex group, and include LOWESS (or LOESS) curves for each sex group.



(b)

Is there evidence from the scatterplot of an association between systolic blood pressure and age after adjusting for sex? Explain your reasoning.

We dont see any association

- The marginal trend is same.
- SBP is also hovering for both sexes in 1 SD of 10 mm/Hg

(c)

Is there evidence from the scatterplot that sex modifies the association between systolic blood pressure and age? Explain your reasoning.

Yes, we do see evidence of sex modifying the association between SBP and age.

- Lines cross.
- There maybe an effect modification, two slopes are not exactly same.

(d)

Perform a statistical analysis to determine if sex modifies the association between systolic blood pressure and age. Provide full statistical inference.

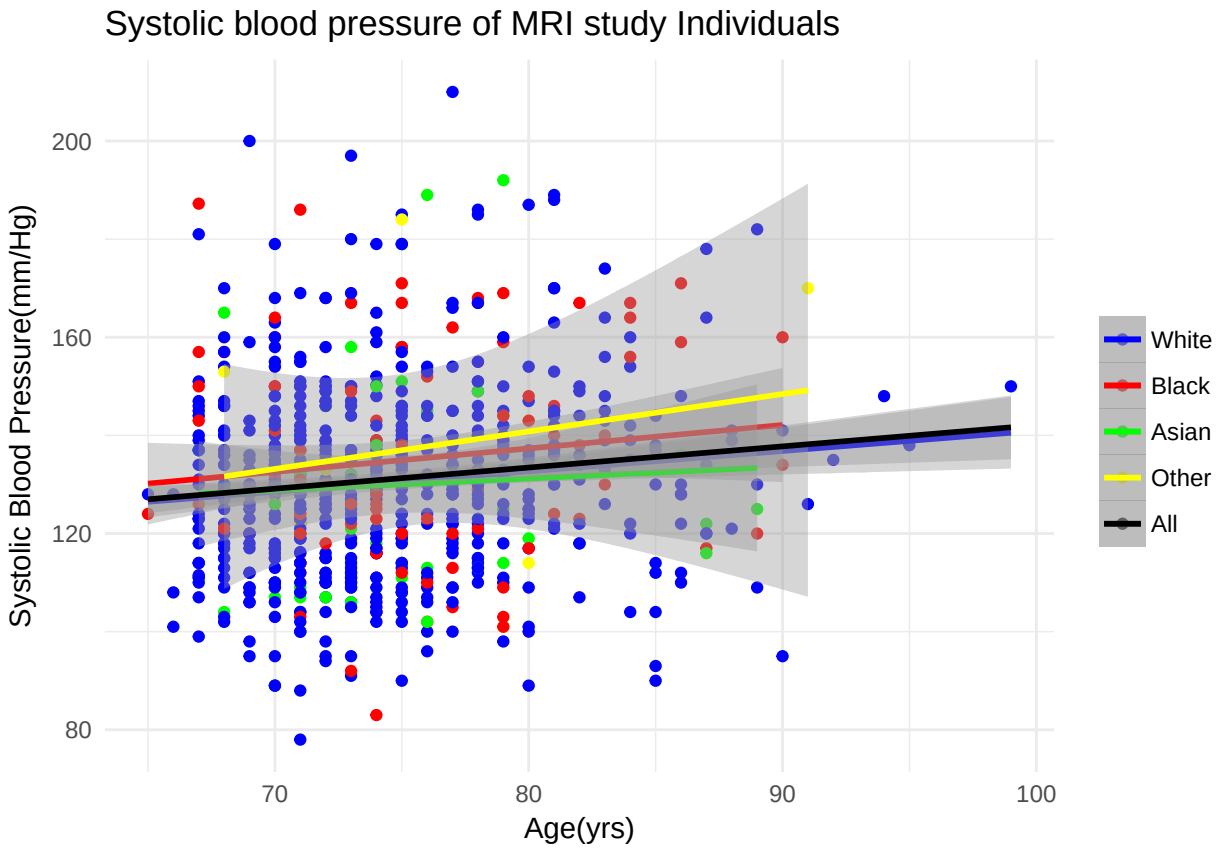
For each 1 year difference in age between two groups that have the same sex, the difference in mean SBP is .5709(95% CI:-1.090 to -0.05) lower in male group. The results are highly atypical of what we might expect if there was no true difference in mean SBP levels between two groups. In summary, we find that there is evidence of sex having an effect on the association between SBP and Age(P-value is found to be 0.03)

Question 2

Now suppose we are interested in examining how mean systolic blood pressure varies by race and age.

(a)

Create a scatterplot of systolic blood pressure versus age. Use different symbols and/or colors for each race group, and include LOWESS (or LOESS) curves for each race group.



(b)

What observations do you make from the scatterplot regarding an association between systolic blood pressure and race.

- 'Whites' inline with the marginal trend (intution says might be skewed on sample)
- 'Other' race having a non-linear trend
- Sampling variability

(c)

Perform a multivariate linear regression analysis with systolic blood pressure as the response and with race and age as predictors. What is the baseline group for race in your regression model. Provide an interpretation of the intercept in your regression model and include the numerical value of the intercept in your interpretation. Is the intercept scientifically useful? Briefly explain.

1= white, 2= black, 3= Asian, 4= other

Baseline group for race is 'white'. The intercept is the estimated mean SBP for white, and the corresponding value is 99.01 mm/Hg. Yes ,it is not scientifically relevant as it gives the default estimated mean SBP for one racial group for newborns.

(d)

Provide an interpretation of the age slope in your regression model in part c, and include the numerical value of the age slope in your interpretation. Is the age slope scientifically useful? Briefly explain.

Age slope is the mean estimated difference in SBP in one year change in age by keeping the race constant. Here the age slope is 0.4213, and it is relevant as it gives the estimated difference in SBP for different racial groups.

(e)

Is race a confounder, precision variable, or neither for the association between systolic blood pressure and age? Explain and provide evidence to support your reasoning.

Race looks to be a precision variable.

- As it reduces the variance when race is added, the confidence intervel becomes tighter(evident in scatter plot)

(f)

Perform a statistical analysis using the multivariate regression model in part c to determine if race is associated with systolic blood pressure after adjusting for age. Provide full statistical inference.

Based on ANOVA, we find that race is not associated with SBP after adjusting for age (P-value is .135), our full model included age and race, and our null model included only age.

Question 3

Perform a multivariate linear regression analysis with systolic blood pressure as the response and with race, sex, age, and an interaction for sex and age as predictors.

(a)

What is the baseline group for race in your regression model. Provide an interpretation of the intercept in your regression model and include the numerical value of the intercept in your interpretation. What, if any, scientific use would you make of the intercept?

Baseline group for race is 'white'. Intercept here is the estimated mean SBP for white newborn females. The numerical value is 76.36(mm/Hg). Scientifically not relevant as it is not what was collected or the subjects were.

(b)

Provide an interpretation of the sex slope in your regression model, and include the numerical value of the sex slope in your interpretation. Is the sex slope scientifically useful? Briefly explain.

Sex slope is the mean estimated difference in SBP between males and females by keeping the race, and age(=0) constant. Here the sex slope is 42.29 higher than female, and it is not relevant as it gives the SBP for newborns .

(c)

Provide an interpretation of the age slope in your regression model, and include the numerical value of the age slope in your interpretation. Is the age slope scientifically useful? Briefly explain.

Age slope is the mean estimated difference in SBP for each 1 yr difference in age by keeping the race, and sex(0=female) constant. Here the age slope is .7302, and it is relevant as it gives the estimated difference in SBP for two female age groups who differ by one year.

(d)

Perform a statistical analysis using the multivariate regression model to determine if age is associated with systolic blood pressure. Provide full statistical inference.

We did ANOVA with null model excluding age, and its interaction with sex, and full model as age,race,sex and sex-age interaction, and found that age is associated with SBP. The associastion is found to be significant at P-value < 0.001.

Optionally, for each year difference in age between two groups that have the same sex and race, the difference in mean SBP is .7302(95% CI:.373 to 1.087), with the older group having a higher SBP. The results are highly atypical of what we might expect if there was no true difference in mean SBP levels between age groups that have the same sex and race.(P-value is 0.0001)

(e)

Perform a statistical analysis using the multivariate regression model to determine if sex is associated with systolic blood pressure. Provide full statistical inference.

Between two sex groups that have the same age and race, the difference in mean SBP is 42.29(95% CI:4.284 to 80.29), with the males having a higher SBP The results are highly atypical of what we might expect if there was no true difference in mean SBP levels between two sex groups that have the same age and race(P-value is .02)

But, doing ANOVA we find that sex is associated with SBP (p-value is 0.0862)

(f)

Perform a statistical analysis using the multivariate regression model to determine if race is associated with systolic blood pressure. Provide full statistical inference.

Doing ANOVA, we find that race is not associated with SBP (P-value is 0.1289), in the ANOVA test, the null-model is with age, sex and its interaction, and full model being age,race,sex and sex-age interaction.

Additionally, doing linear regression, we find that between the racial groups 'white' and 'black' that have the same age and sex, the difference in mean SBP is 4.29(95% CI:0.03 to 8.55), with the 'black' group having a higher SBP. The results are highly atypical of what we might expect if there was no true difference in mean SBP levels between two racial groups('white' and 'black') that have the same age and sex (P-value is .04).

(g)

Perform a statistical analysis using the multivariate regression model for testing the null hypothesis that both age and sex are not associated with systolic blood pressure. Provide full statistical inference.

Doing ANOVA by excluding the sex and age interaction in null model and as age,race,sex and sex-age interactionin main model, we see that both age and sex are associated with systolic blood pressure (P-value is found to be < 0.01) i.e. we reject the null hypothesis.