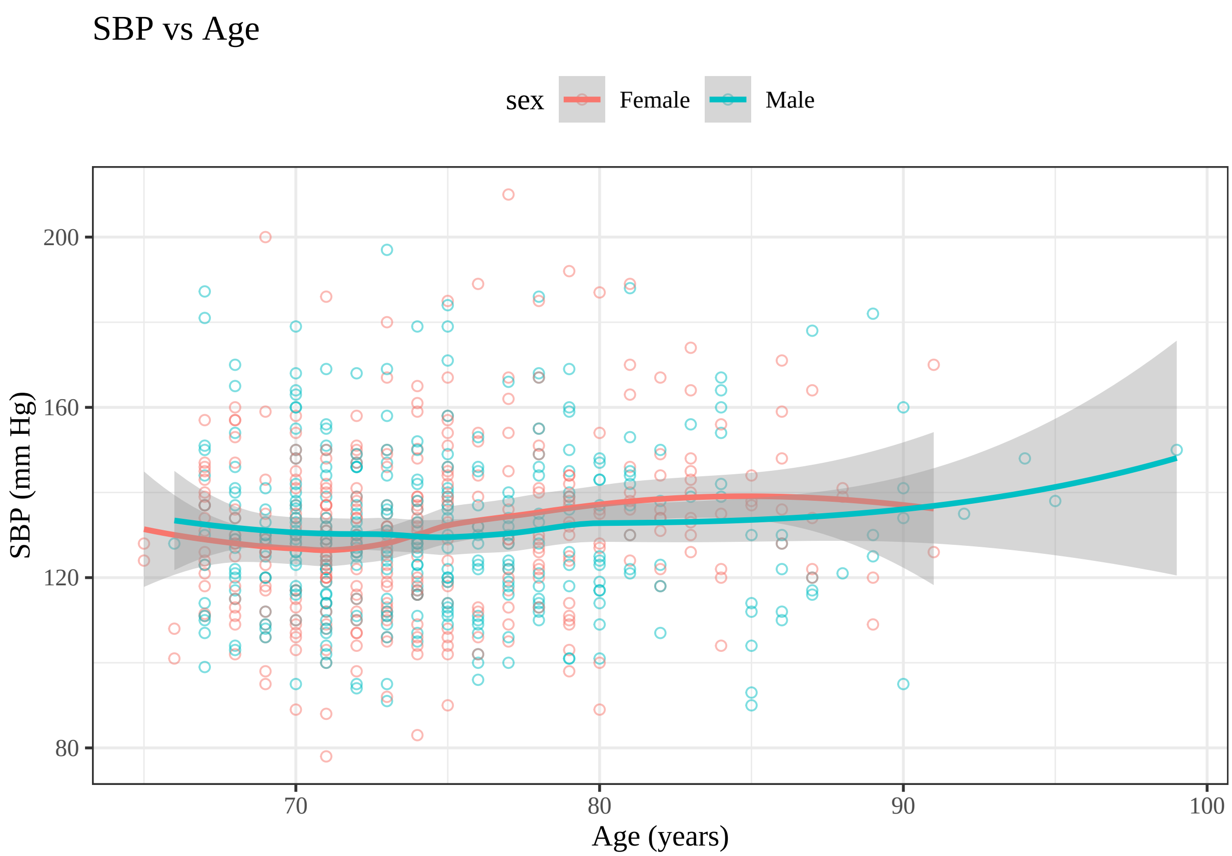
Homework 05

Spencer Pease

2/14/2020

(Q1) Systolic blood pressure by age and sex

(Q1.a)



(Q1.b)

From the scatter plot, fit with loess smooths for each sex, there does not appear to be any conclusive evidence that sex affects the association between SBP and age. From the scatter plot itself, we see that there is no clear separation along either axis by sex, and while the loess smooths have different patterns, they follow each other closely enough to fall within the standard error margin of each other. The upward trend in SBP for 90+ year-old males is likely due to the small availability of data in that subset of the population, leading to a sample without any female representation.

(Q1.c)

Table 1: Inferential statistics of SBP vs age, adjusted for sex

Parameter	Estimate	Robust SE	95%L	95%H	Pr(>|t|)
(Intercept)	78.035	14.276	50.002	106.067	6.55e-08
age	0.716	0.192	0.338	1.093	2.12e-04
male	41.930	20.493	1.690	82.169	4.11e-02
age:male	-0.568	0.275	-1.108	-0.028	3.91e-02

In the linear model of SBP vs age, adjusted for sex, fit using robust standard errors, we find that there is a statistically significant first-order relationship between sbp and age and the interaction between age and sex. For two groups of females differing by one year in age, mean sbp is 0.72 mm Hg higher in the older group (95% CI: 0.34, 1.09). For two groups of males differing by one year in age, mean sbp is 0.15 mm Hg higher in the older group.

(Q1.d)

This model show there is a difference between the association of sbp and age when accounting for sex, but that difference is slight and both are trending in the same direction. Looking at the scatter plot confirms this: in the range of data where there are both male and female observations, the net change in slope of the loess line for females is greater than the net change for males.

(Q2) Systolic blood pressure by age and race

(Q2.a)

Model of race-adjusted association between mean systolic blood pressure and age:

$$E(sbp \mid age, race) = \beta_0 + \beta_1 \cdot age + \beta_{2-4} \cdot race + \beta_{5-7} \cdot age \cdot race$$

where *race* is treated as a categorical variable, with possible values *White*, *Black*, *Asian*, and *Other* (treated as the reference category).

(Q2.b)

Table 2: Inferential statistics of SBP vs age, adjusted for race

Parameter	df	Estimate	Robust SE	95%L	95%H	Pr(>F)
Intercept	1	79.702	63.587	-45.159	204.563	0.210
age	1	0.764	0.878	-0.960	2.487	0.385
race	3	-	-	-	-	0.978
| race-White	1	21.198	64.624	-105.699	148.094	0.743
| race-Black	1	21.859	69.196	-114.015	157.733	0.752
| race-Asian	1	31.474	72.397	-110.685	173.633	0.664
age:race	3	-	-	-	-	0.961
| age:race-White	1	-0.369	0.891	-2.119	1.381	0.679

Parameter	df	Estimate	Robust SE	95%L	95%H	Pr(>F)
| age:race-Black	1	-0.316	0.950	-2.181	1.548	0.739
| age:race-Asian	1	-0.513	0.991	-2.459	1.433	0.605

In this linear model of the association between sbp and age, adjusted for race and fit with robust standard errors, we do not find significant evidence that there is a first-order trend in our sample indicating different race-adjusted age patterns for sbp (P-value .96). In this model, race is treated as a categorical variable, with *other* being the baseline race. A 3-degree-of-freedom test was performed, showing that any race of *White*, *Black*, or *Asian* was significantly different from the base case in regarding any interaction with age.

(Q3) Systolic blood pressure by age, sex, and race

(Q3.a)

Model of the association between SBP and age, adjusted for race and sex:

$$E(sbp \mid age, male, race) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot male + \beta_{3-5} \cdot race + (\beta_6 \cdot male + \beta_{7-9} \cdot race) \cdot age$$

Table 3: Inferential statistics of SBP vs age, adjusted for race and sex

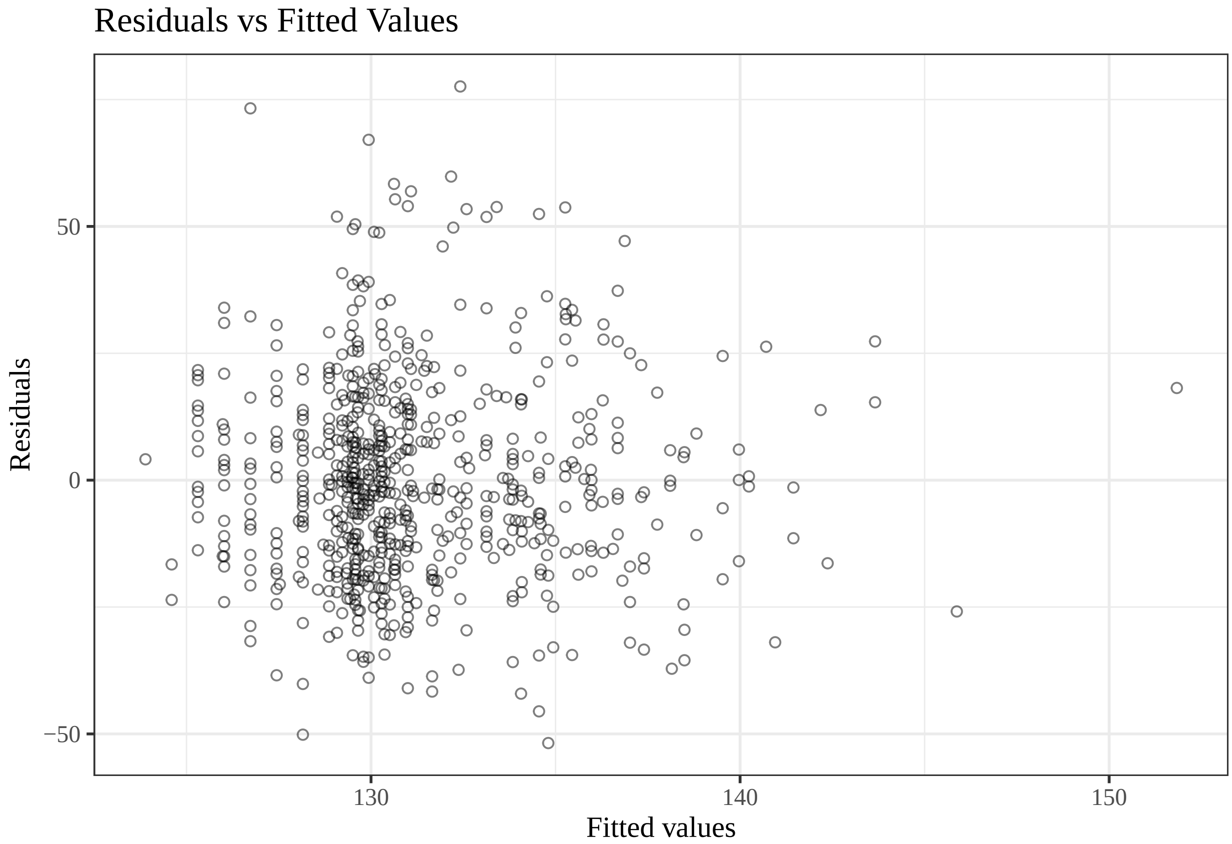
Parameter	df	Estimate	Robust SE	95%L	95%H	Pr(>F)
Intercept	1	71.156	58.404	-43.527	185.839	0.224
age	1	0.887	0.803	-0.690	2.463	0.270
male	1	41.799	20.271	1.995	81.602	0.040
race	3	-	-	-	-	0.982
| race-White	1	6.521	59.823	-110.949	123.991	0.913
| race-Black	1	9.040	64.759	-118.124	136.203	0.889
| race-Asian	1	20.285	67.352	-111.968	152.538	0.763
age:male	1	-0.568	0.272	-1.101	-0.034	0.037
age:race	3	-	-	-	-	0.970
| age:race-White	1	-0.176	0.822	-1.789	1.438	0.831
| age:race-Black	1	-0.149	0.885	-1.887	1.590	0.867
| age:race-Asian	1	-0.371	0.918	-2.174	1.432	0.686

(Q3.b)

Table 4: 95% PI for SBP among 70-year-old black women

Prediction	95%L	95%H
131.851	92.254	171.448

(Q3.c)



The residual plot shows that, for a given fitted value, the residuals tend to lie symmetrically around zero. This confirms that variance across groups in our sample is homoscedastic, which is one of the underlying assumptions required for prediction.

(Q4) Odds ratios

(Q4.a-c)

Odds ratio for disease, given exposure:

$$OR_{(D|E=1)} = \frac{Odds_{D|E=1}}{Odds_{D|E=0}} = \frac{P_{(D|E=1)} \div (1 - P_{(D|E=1)})}{P_{(D|E=0)} \div (1 - P_{(D|E=0)})} = \frac{d/c}{b/a}$$

Odds ratio for exposure, given disease:

$$OR_{(E|D=1)} = \frac{Odds_{E|D=1}}{Odds_{E|D=0}} = \frac{P_{(E|D=1)} \div (1 - P_{(E|D=1)})}{P_{(E|D=0)} \div (1 - P_{(E|D=0)})} = \frac{d/b}{c/a}$$

Comparing these two odds ratios, we see they are identical:

$$OR_{(D|E=1)} = \frac{d/c}{b/a} = \frac{ad}{bc} = \frac{d/b}{c/a} = OR_{(E|D=1)}$$