

MATH 423 Final Project

Jeremy Kim - 260620525

2016-12-20

bpdia

```
bpdia.fit1<-lm(bpdia ~ age + alcdays + alcyer + bmi + dirchol + educ + gender + hhinc + homeown + homerooms + married + physact + poverty + pulse + race3 + totchol + weight)
```

```
# remove some effects
bpdia.fit2<-step(bpdia.fit1,test='F')
```

```
anova(bpdia.fit1,bpdia.fit2,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpdia ~ age + alcdays + alcyer + bmi + dirchol + educ + gender +
##      hhinc + homeown + homerooms + married + physact + poverty +
##      pulse + race3 + totchol + weight
## Model 2: bpdia ~ age + gender + hhinc + married + poverty + pulse + race3 +
##      totchol + weight
##   Res.Df    RSS  Df Sum of Sq    F Pr(>F)
## 1    2064 235300
## 2    2076 236128 -12   -827.86 0.6052 0.8395
```

Has high F value, thus predictors removed from fit 1 are not significant enough. Fit 2 is adequate.

Now add some second order interactions.

```
# add some second order interactions
bpdia.fit3<-step(bpdia.fit2,test='F', scope= list(upper=update(bpdia.fit2, ~.*.), lower= bpdia.fit2))
```

```
anova(bpdia.fit2,bpdia.fit3, test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpdia ~ age + gender + hhinc + married + poverty + pulse + race3 +
##      totchol + weight
## Model 2: bpdia ~ age + gender + hhinc + married + poverty + pulse + race3 +
##      totchol + weight + age:married + hhinc:pulse + pulse:totchol +
##      gender:race3 + married:totchol + age:totchol + gender:pulse +
##      age:hhinc + age:poverty + age:race3 + race3:totchol + hhinc:race3 +
##      hhinc:totchol + married:race3 + hhinc:poverty + gender:hhinc +
##      age:gender + gender:totchol + gender:poverty
##   Res.Df    RSS  Df Sum of Sq    F    Pr(>F)
## 1    2076 236128
## 2    1914 185732 162    50396 3.2058 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# low p value fit 3 is adequate
```

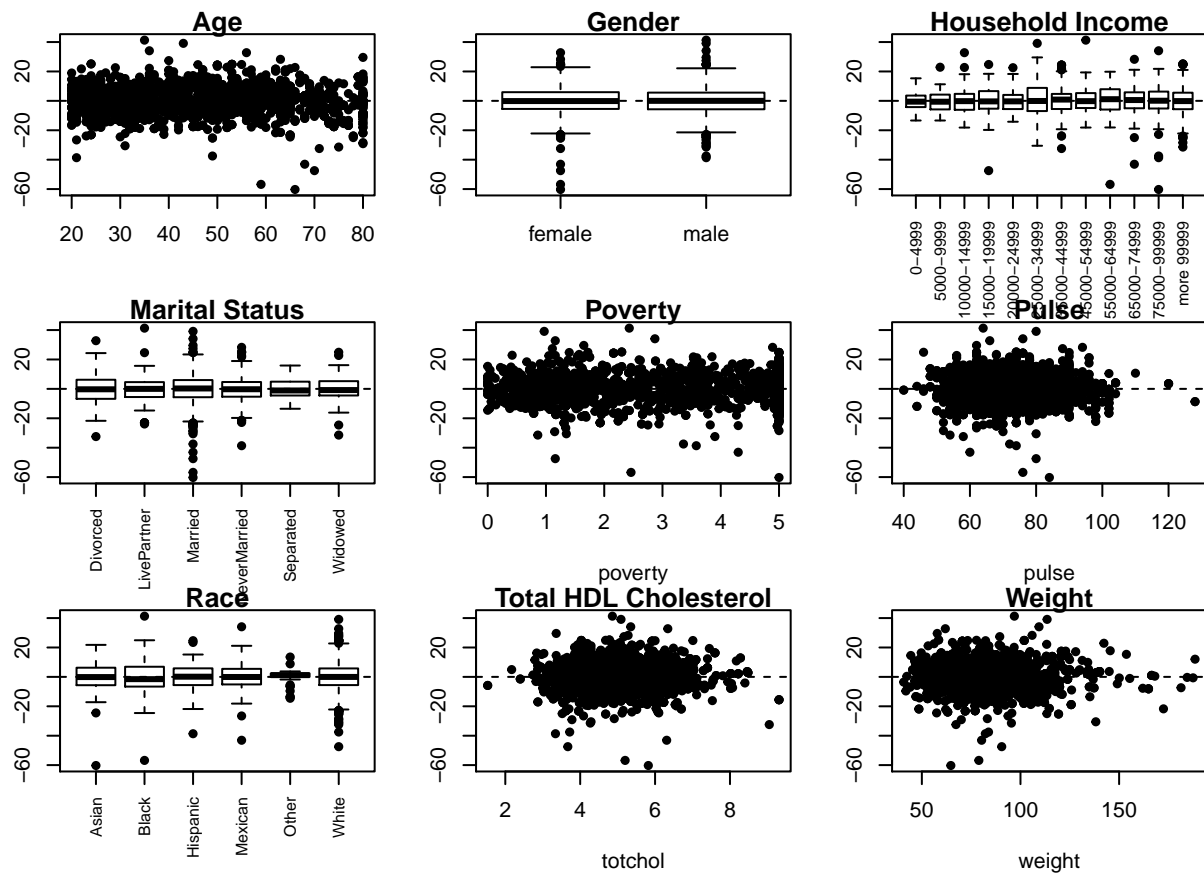
Has low F value, thus interactions added are significant enough. Fit 3 is adequate.

Model adequacy for bpdia

##	Rsqr	Adj.Rsqr	Cp	AIC	BIC
## 1	0.1337	0.1174	400.8060	15977.50	16209.22
## 2	0.1307	0.1194	385.3373	15960.89	16124.79
## 3	0.3162	0.2487	200.0000	15779.79	16859.25

Fit 3 has lowest AIC and highest R_{adj}^2 , thus it is most adequate.

Residual plots for bpdia



Residual plots seem reasonable with constant variance. Factor predictors have less consistent residuals but they seem to even out.

Diastolic blood pressure is higher for males (7.684), slightly lower for older people (-.4278), higher for less poor people (29.171, though offsetted by the interaction between poverty index and hhinc), lower for those with higher cholesterol (-7.197), highest for middle class (45k-75k hhinc) and lowest for 25k-45k hhinc, highest for other races and lowest for Blacks and Hispanics.

bpsys

```
bpsys.fit1<-lm(bpsys ~ age + alcd day + alcy year + bmi + dirchol + educ + gender + hhinc + homeown + homer
```

```
bpsys.fit2<-step(bpsys.fit1,test='F')
```

```
anova(bpsys.fit1,bpsys.fit2,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpsys ~ age + alcd day + alcy year + bmi + dirchol + educ + gender +
##      hhinc + homeown + homerooms + married + physact + poverty +
##      pulse + race3 + totchol + weight
## Model 2: bpsys ~ age + alcd day + alcy year + bmi + dirchol + educ + gender +
##      hhinc + homeown + married + physact + poverty + pulse + race3 +
##      totchol
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1    2064 384551
## 2    2066 385184 -2    -632.39 1.6971 0.1835
```

Has high F value, thus predictors removed from fit 1 are not significant enough. Fit 2 is adequate.

Let's see if there are more predictors that can be dropped

```
drop1(bpsys.fit2, test='F')
```

```
## Single term deletions
##
## Model:
## bpsys ~ age + alcd day + alcy year + bmi + dirchol + educ + gender +
##      hhinc + homeown + married + physact + poverty + pulse + race3 +
##      totchol
##           Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                385184 11038
## age          1      54608 439792 11314 292.8994 < 2.2e-16 ***
## alcd day      1       2018 387202 11047  10.8243  0.001019 **
## alcy year     1       1977 387161 11046  10.6025  0.001148 **
## bmi           1       7665 392849 11077  41.1149 1.774e-10 ***
## dirchol       1        492 385676 11038   2.6400  0.104354
## educ          4       1922 387105 11040   2.5766  0.035880 *
## gender        1       6114 391298 11069  32.7957 1.174e-08 ***
## hhinc         11       7294 392478 11055   3.5568 5.632e-05 ***
## homeown        2       2239 387423 11046   6.0046  0.002511 **
## married        5       4401 389585 11052   4.7214  0.000272 ***
## physact        1        456 385640 11038   2.4468  0.117917
## poverty        1       4737 389921 11061  25.4102 5.038e-07 ***
## pulse          1       1838 387022 11046   9.8602  0.001713 **
## race3          5       2132 387315 11039   2.2868  0.043821 *
## totchol        1       4900 390084 11062  26.2843 3.222e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
bpsys.fit3<-update(bpsys.fit2, ~. - dirchol - physact)
anova(bpsys.fit2,bpsys.fit3,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpsys ~ age + alcd day + alcy year + bmi + dirchol + educ + gender +
##      hhinc + homeown + married + physact + poverty + pulse + race3 +
##      totchol
## Model 2: bpsys ~ age + alcd day + alcy year + bmi + educ + gender + hhinc +
##      homeown + married + poverty + pulse + race3 + totchol
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1    2066 385184
## 2    2068 386168 -2    -983.89 2.6386 0.0717 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Has high F value, thus predictors removed from fit 2 are not significant enough. Fit 3 is adequate. Now add some second order interactions.

```
# add some second order interactions
add1(bpsys.fit3,test='F', scope= update(bpsys.fit3, ~.*.))
```

```
## Single term additions
##
## Model:
## bpsys ~ age + alcd day + alcy year + bmi + educ + gender + hhinc +
##      homeown + married + poverty + pulse + race3 + totchol
##           Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                386168 11039
## age:alcd day         1      704.8 385463 11037  3.7792 0.0520281 .
## age:alcy year         1      117.2 386050 11040  0.6276 0.4283231
## age:bmi               1     1136.3 385031 11035  6.1001 0.0135975 *
## age:educ              4      929.5 385238 11042  1.2450 0.2897154
## age:gender            1     7558.4 378609 10999 41.2649 1.645e-10 ***
## age:hhinc            11     3035.6 383132 11044  1.4816 0.1314237
## age:homeown           2     1524.4 384643 11035  4.0940 0.0168076 *
## age:married           5     1281.8 384886 11042  1.3741 0.2309597
## age:poverty           1     1619.2 384548 11032  8.7036 0.0032116 **
## age:pulse             1         0.3 386167 11041  0.0014 0.9696938
## age:race3             5     3306.8 382861 11031  3.5637 0.0032698 **
## age:totchol           1     3947.7 382220 11019 21.3486 4.065e-06 ***
## alcd day:alcy year     1       243.2 385924 11040  1.3024 0.2538989
## alcd day:bmi           1     1625.6 384542 11032  8.7377 0.0031522 **
## alcd day:educ          4     6006.4 380161 11014  8.1527 1.606e-06 ***
## alcd day:gender        1         0.3 386167 11041  0.0016 0.9684540
## alcd day:hhinc        11     8982.7 377185 11011  4.4534 1.153e-06 ***
## alcd day:homeown       2     1117.9 385050 11037  2.9990 0.0500545 .
## alcd day:married       5     1550.9 384617 11040  1.6638 0.1400526
## alcd day:poverty       1     2150.7 384017 11029 11.5766 0.0006806 ***
## alcd day:pulse         1     1814.5 384353 11031  9.7583 0.0018099 **
## alcd day:race3         5     6408.7 379759 11014  6.9629 1.851e-06 ***
## alcd day:totchol       1         19.9 386148 11041  0.1066 0.7440710
```

## alcyear:bmi	1	1398.4	384769	11033	7.5121	0.0061815	**
## alcyear:educ	4	1465.1	384703	11039	1.9651	0.0972550	.
## alcyear:gender	1	3.1	386165	11041	0.0165	0.8979136	
## alcyear:hhinc	11	8616.5	377551	11014	4.2677	2.618e-06	***
## alcyear:homeown	2	2022.1	384146	11032	5.4375	0.0044127	**
## alcyear:married	5	8929.1	377239	11000	9.7661	3.090e-09	***
## alcyear:poverty	1	412.4	385755	11039	2.2098	0.1372865	
## alcyear:pulse	1	428.0	385740	11039	2.2932	0.1300940	
## alcyear:race3	5	3427.3	382740	11030	3.6947	0.0024799	**
## alcyear:totchol	1	3354.5	382813	11023	18.1126	2.175e-05	***
## bmi:educ	4	1124.4	385043	11041	1.5069	0.1975439	
## bmi:gender	1	111.9	386056	11040	0.5994	0.4389011	
## bmi:hhinc	11	5117.3	381050	11033	2.5113	0.0038586	**
## bmi:homeown	2	32.3	386135	11043	0.0864	0.9172348	
## bmi:married	5	2984.4	383183	11033	3.2136	0.0067951	**
## bmi:poverty	1	107.3	386060	11040	0.5743	0.4486602	
## bmi:pulse	1	151.3	386016	11040	0.8104	0.3680994	
## bmi:race3	5	1900.4	384267	11039	2.0405	0.0701666	.
## bmi:totchol	1	24.5	386143	11041	0.1311	0.7173212	
## educ:gender	4	862.7	385305	11042	1.1553	0.3287438	
## educ:hhinc	42	14580.0	371588	11042	1.8927	0.0005103	***
## educ:homeown	8	3700.5	382467	11035	2.4914	0.0108625	*
## educ:married	20	8965.6	377202	11030	2.4339	0.0003836	***
## educ:poverty	4	2261.9	383906	11035	3.0402	0.0164034	*
## educ:pulse	4	1202.3	384965	11040	1.6115	0.1686680	
## educ:race3	20	5067.0	381101	11051	1.3615	0.1305794	
## educ:totchol	4	398.9	385769	11045	0.5335	0.7111274	
## gender:hhinc	11	4918.8	381249	11034	2.4127	0.0056039	**
## gender:homeown	2	1442.3	384725	11035	3.8725	0.0209565	*
## gender:married	5	3327.0	382841	11031	3.5856	0.0031222	**
## gender:poverty	1	35.4	386132	11041	0.1893	0.6635498	
## gender:pulse	1	567.4	385600	11038	3.0414	0.0813126	.
## gender:race3	5	905.3	385262	11044	0.9695	0.4350297	
## gender:totchol	1	65.1	386103	11041	0.3483	0.5551184	
## hhinc:homeown	20	4895.1	381273	11052	1.3147	0.1580006	
## hhinc:married	55	19068.7	367099	11042	1.9012	8.826e-05	***
## hhinc:poverty	11	3707.1	382461	11041	1.8125	0.0469411	*
## hhinc:pulse	11	3639.4	382528	11041	1.7791	0.0523669	.
## hhinc:race3	53	15267.8	370900	11060	1.5650	0.0060827	**
## hhinc:totchol	11	2752.3	383415	11046	1.3423	0.1944641	
## homeown:married	9	6085.1	380083	11024	3.6627	0.0001471	***
## homeown:poverty	2	191.0	385977	11042	0.5112	0.5998624	
## homeown:pulse	2	212.2	385955	11042	0.5680	0.5667612	
## homeown:race3	9	3715.5	382452	11037	2.2226	0.0182833	*
## homeown:totchol	2	1088.0	385080	11037	2.9187	0.0542284	.
## married:poverty	5	1649.8	384518	11040	1.7703	0.1156682	
## married:pulse	5	850.7	385317	11044	0.9109	0.4728537	
## married:race3	22	9111.8	377056	11033	2.2474	0.0007765	***
## married:totchol	5	5002.0	381166	11022	5.4145	5.923e-05	***
## poverty:pulse	1	123.7	386044	11040	0.6626	0.4157421	
## poverty:race3	5	1333.0	384835	11042	1.4292	0.2105323	
## poverty:totchol	1	226.2	385941	11040	1.2113	0.2711973	
## pulse:race3	5	2270.9	383897	11036	2.4407	0.0324479	*
## pulse:totchol	1	134.6	386033	11040	0.7205	0.3960673	

```
## race3:totchol    5    2539.2 383628 11035  2.7309 0.0182150 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
bpsys.fit4<-update(bpsys.fit3, ~. +age:gender+age:poverty + age:race3 + age:totchol + alcd:
anova(bpsys.fit4,bpsys.fit3,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpsys ~ age + alcd: + alcd:year + bmi + educ + gender + hhinc +
##   homeown + married + poverty + pulse + race3 + totchol + age:gender +
##   age:poverty + age:race3 + age:totchol + alcd:bmi + alcd:educ +
##   alcd:hhinc + alcd:poverty + alcd:pulse + alcd:race3 +
##   educ:hhinc + educ:married + gender:hhinc + gender:married +
##   hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##   married:totchol
## Model 2: bpsys ~ age + alcd: + alcd:year + bmi + educ + gender + hhinc +
##   homeown + married + poverty + pulse + race3 + totchol
##   Res.Df    RSS    Df Sum of Sq    F    Pr(>F)
## 1    1816 285231
## 2    2068 386168 -252    -100937 2.5502 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Has low F value, thus interactions added are significant enough. Fit 4 is adequate.

```
# drop some variable
drop1(bpsys.fit4, test='F')
```

```
## Single term deletions
##
## Model:
## bpsys ~ age + alcd: + alcd:year + bmi + educ + gender + hhinc +
##   homeown + married + poverty + pulse + race3 + totchol + age:gender +
##   age:poverty + age:race3 + age:totchol + alcd:bmi + alcd:educ +
##   alcd:hhinc + alcd:poverty + alcd:pulse + alcd:race3 +
##   educ:hhinc + educ:married + gender:hhinc + gender:married +
##   hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##   married:totchol
##           Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                285231 10906
## alcd:year            1    1513.8 286745 10915  9.6380 0.0019354 **
## age:gender           1    1556.2 286787 10915  9.9077 0.0016726 **
## age:poverty          1     563.6 285795 10908  3.5886 0.0583363 .
## age:race3            5     384.6 285616 10898  0.4897 0.7841408
## age:totchol          1    2303.8 287535 10920 14.6676 0.0001326 ***
## alcd:bmi             1     394.4 285625 10906  2.5113 0.1132054
## alcd:educ            4    1946.2 287177 10912  3.0977 0.0148946 *
## alcd:hhinc          11    5543.3 290774 10924  3.2084 0.0002440 ***
## alcd:poverty         1     213.1 285444 10905  1.3568 0.2442496
## alcd:pulse           1      57.1 285288 10904  0.3637 0.5465172
## alcd:race3           5     668.0 285899 10900  0.8506 0.5138301
## educ:hhinc          42   12919.8 298151 10915  1.9585 0.0002598 ***
```

```
## educ:married      20      6559.0 291790 10913  2.0880 0.0032201 **
## gender:hhinc      11      6682.0 291913 10932  3.8676 1.522e-05 ***
## gender:married     5      1733.1 286964 10908  2.2069 0.0511823 .
## hhinc:married     54     14858.7 300090 10904  1.7519 0.0006814 ***
## hhinc:race3       52     10831.1 296062 10880  1.3261 0.0611020 .
## homeown:married    9      5197.6 290429 10926  3.6769 0.0001414 ***
## married:race3     22      7546.2 292777 10916  2.1839 0.0011953 **
## married:totchol    5      2480.1 287711 10914  3.1581 0.0076424 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
bpsys.fit5<-update(bpsys.fit4, ~. - age:race3 - alcd day:bmi - alcd day:poverty - alcd day:pulse - alcd day:race3
anova(bpsys.fit5,bpsys.fit4,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpsys ~ age + alcd day + alcd year + bmi + educ + gender + hhinc +
##      homeown + married + poverty + pulse + race3 + totchol + age:gender +
##      age:poverty + age:totchol + alcd day:educ + alcd day:hhinc +
##      educ:hhinc + educ:married + gender:hhinc + gender:married +
##      hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##      married:totchol
## Model 2: bpsys ~ age + alcd day + alcd year + bmi + educ + gender + hhinc +
##      homeown + married + poverty + pulse + race3 + totchol + age:gender +
##      age:poverty + age:race3 + age:totchol + alcd day:bmi + alcd day:educ +
##      alcd day:hhinc + alcd day:poverty + alcd day:pulse + alcd day:race3 +
##      educ:hhinc + educ:married + gender:hhinc + gender:married +
##      hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##      married:totchol
##      Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1      1829 287490
## 2      1816 285231 13      2258.9 1.1063 0.3484
```

Has high F value, thus predictors removed are not significant. Fit 5 is adequate.

```
# add some
add1(bpsys.fit5, test="F", scope = update(bpsys.fit3, ~.*.))
```

```
## Single term additions
##
## Model:
## bpsys ~ age + alcd day + alcd year + bmi + educ + gender + hhinc +
##      homeown + married + poverty + pulse + race3 + totchol + age:gender +
##      age:poverty + age:totchol + alcd day:educ + alcd day:hhinc +
##      educ:hhinc + educ:married + gender:hhinc + gender:married +
##      hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##      married:totchol
##      Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                287490 10896
## age:alcd day          1      206.4 287283 10897  1.3134 0.2519343
## age:alcd year          1       30.7 287459 10898  0.1955 0.6584273
## age:bmi                1     1091.9 286398 10890  6.9695 0.0083617 **
## age:educ              4       347.6 287142 10902  0.5522 0.6974127
```

## age:hhinc	11	1719.8	285770	10906	0.9946	0.4487183	
## age:homeown	2	819.2	286671	10894	2.6105	0.0737710	.
## age:married	5	1460.5	286029	10895	1.8628	0.0977606	.
## age:pulse	1	574.9	286915	10894	3.6628	0.0557971	.
## age:race3	5	819.6	286670	10900	1.0430	0.3905675	
## alcd day:alcy year	1	3.4	287486	10898	0.0214	0.8836475	
## alcd day:bmi	1	643.3	286847	10893	4.0996	0.0430395	*
## alcd day:gender	1	142.7	287347	10897	0.9077	0.3408406	
## alcd day:homeown	2	30.6	287459	10900	0.0972	0.9074019	
## alcd day:married	5	842.3	286647	10900	1.0720	0.3739282	
## alcd day:poverty	1	93.3	287397	10897	0.5931	0.4413078	
## alcd day:pulse	1	70.1	287420	10898	0.4458	0.5043974	
## alcd day:race3	5	1361.7	286128	10896	1.7361	0.1230993	
## alcd day:tot chol	1	533.1	286957	10894	3.3959	0.0655198	.
## alcy year:bmi	1	1340.2	286150	10888	8.5618	0.0034752	**
## alcy year:educ	4	1054.8	286435	10896	1.6802	0.1519086	
## alcy year:gender	1	164.5	287325	10897	1.0465	0.3064484	
## alcy year:hhinc	11	5234.1	282256	10879	3.0648	0.0004387	***
## alcy year:homeown	2	1190.0	286300	10891	3.7969	0.0226169	*
## alcy year:married	5	5909.9	281580	10862	7.6566	3.919e-07	***
## alcy year:poverty	1	52.7	287437	10898	0.3353	0.5625994	
## alcy year:pulse	1	420.3	287070	10895	2.6764	0.1020155	
## alcy year:race3	5	1205.7	286284	10897	1.5363	0.1752740	
## alcy year:tot chol	1	1926.9	285563	10884	12.3347	0.0004554	***
## bmi:educ	4	760.8	286729	10898	1.2106	0.3042458	
## bmi:gender	1	65.2	287425	10898	0.4148	0.5196335	
## bmi:hhinc	11	2408.5	285081	10900	1.3963	0.1677615	
## bmi:homeown	2	152.5	287337	10899	0.4849	0.6158521	
## bmi:married	5	1377.0	286113	10896	1.7557	0.1188232	
## bmi:poverty	1	439.0	287051	10895	2.7957	0.0946908	.
## bmi:pulse	1	135.0	287355	10897	0.8589	0.3541580	
## bmi:race3	5	709.3	286781	10901	0.9023	0.4785964	
## bmi:tot chol	1	275.8	287214	10896	1.7551	0.1854036	
## educ:gender	4	154.6	287335	10903	0.2456	0.9124390	
## educ:homeown	8	2054.5	285435	10897	1.6384	0.1090796	
## educ:poverty	4	4318.4	283171	10872	6.9579	1.478e-05	***
## educ:pulse	4	1004.4	286485	10897	1.5996	0.1718128	
## educ:race3	19	3869.5	283620	10906	1.2997	0.1726521	
## educ:tot chol	4	600.2	286890	10900	0.9545	0.4314762	
## gender:homeown	2	451.7	287038	10897	1.4374	0.2378024	
## gender:poverty	1	6.2	287484	10898	0.0394	0.8427467	
## gender:pulse	1	0.6	287489	10898	0.0039	0.9502575	
## gender:race3	5	1339.6	286150	10896	1.7078	0.1295306	
## gender:tot chol	1	12.3	287478	10898	0.0781	0.7799721	
## hhinc:homeown	19	4764.3	282725	10899	1.6053	0.0470554	*
## hhinc:poverty	11	1965.7	285524	10904	1.1378	0.3269696	
## hhinc:pulse	11	2729.5	284760	10898	1.5842	0.0969652	.
## hhinc:tot chol	11	3609.9	283880	10892	2.1017	0.0175081	*
## homeown:poverty	2	794.1	286696	10894	2.5302	0.0799243	.
## homeown:pulse	2	254.3	287236	10898	0.8087	0.4456144	
## homeown:race3	8	3131.4	284358	10889	2.5066	0.0104265	*
## homeown:tot chol	2	1061.6	286428	10892	3.3858	0.0340633	*
## married:poverty	5	1460.0	286030	10895	1.8621	0.0978840	.
## married:pulse	5	1812.6	285677	10893	2.3146	0.0415657	*


```
## poverty:pulse      1      818.2 286672 10892  5.2172 0.0224782 *
## poverty:race3     5     1472.0 286018 10895  1.8774 0.0951601 .
## poverty:totchol   1       15.3 287475 10898  0.0973 0.7550941
## pulse:race3       5     1159.9 286330 10898  1.4778 0.1938663
## pulse:totchol     1     2025.3 285465 10883 12.9692 0.0003251 ***
## race3:totchol     5       785.5 286704 10900  0.9994 0.4165563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
bpsys.fit6<- update(bpsys.fit5, ~. + alcyar:hhinc+alcyar:married +alcyar:totchol + educ:poverty + pu
anova(bpsys.fit5, bpsys.fit6, test='F')
```

```
## Analysis of Variance Table
##
## Model 1: bpsys ~ age + alcdays + alcyar + bmi + educ + gender + hhinc +
##   homeown + married + poverty + pulse + race3 + totchol + age:gender +
##   age:poverty + age:totchol + alcdays:educ + alcdays:hhinc +
##   educ:hhinc + educ:married + gender:hhinc + gender:married +
##   hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##   married:totchol
## Model 2: bpsys ~ age + alcdays + alcyar + bmi + educ + gender + hhinc +
##   homeown + married + poverty + pulse + race3 + totchol + age:gender +
##   age:poverty + age:totchol + alcdays:educ + alcdays:hhinc +
##   educ:hhinc + educ:married + gender:hhinc + gender:married +
##   hhinc:married + hhinc:race3 + homeown:married + married:race3 +
##   married:totchol + alcyar:hhinc + alcyar:married + alcyar:totchol +
##   educ:poverty + pulse:totchol
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1    1829 287490
## 2    1807 268080 22    19410 5.947 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

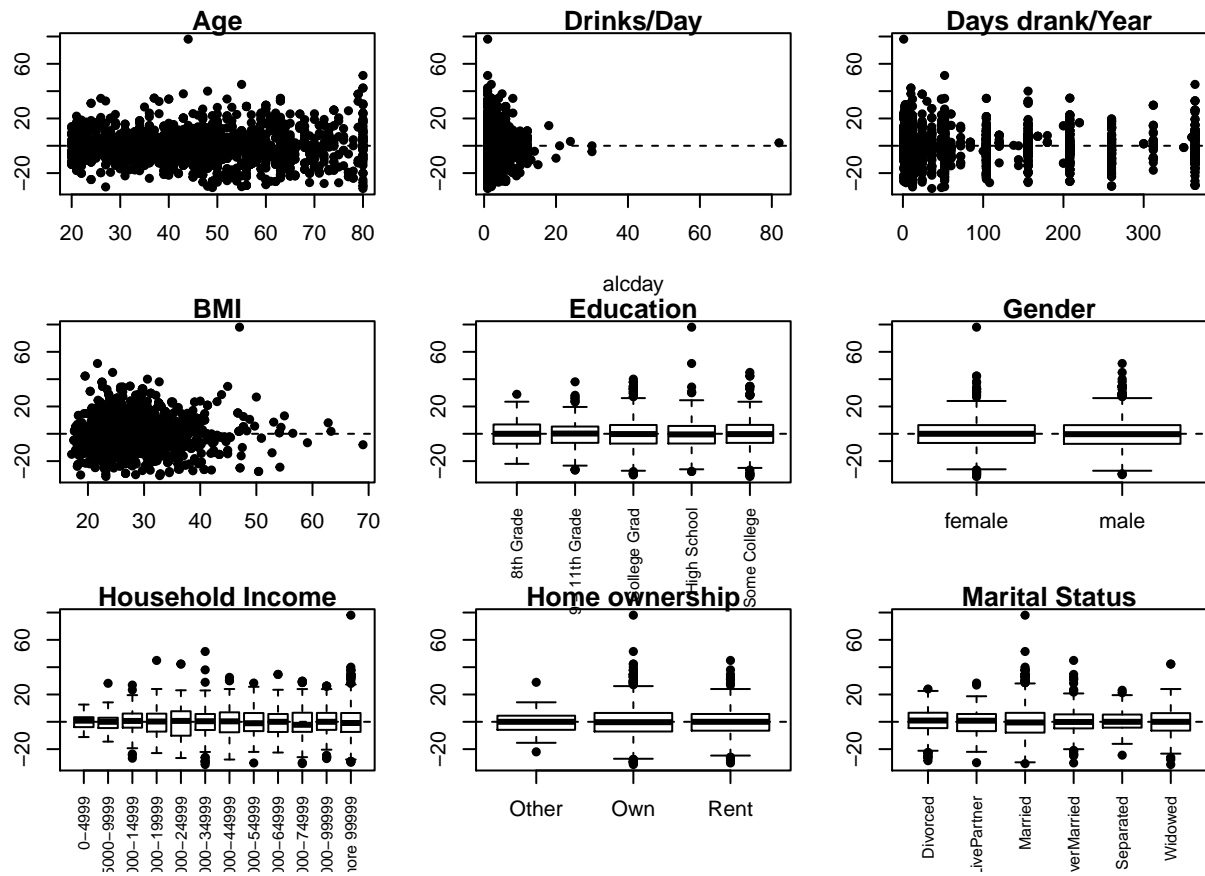
Has low F value, thus predictors added are significant enough. Fit 6 is adequate.

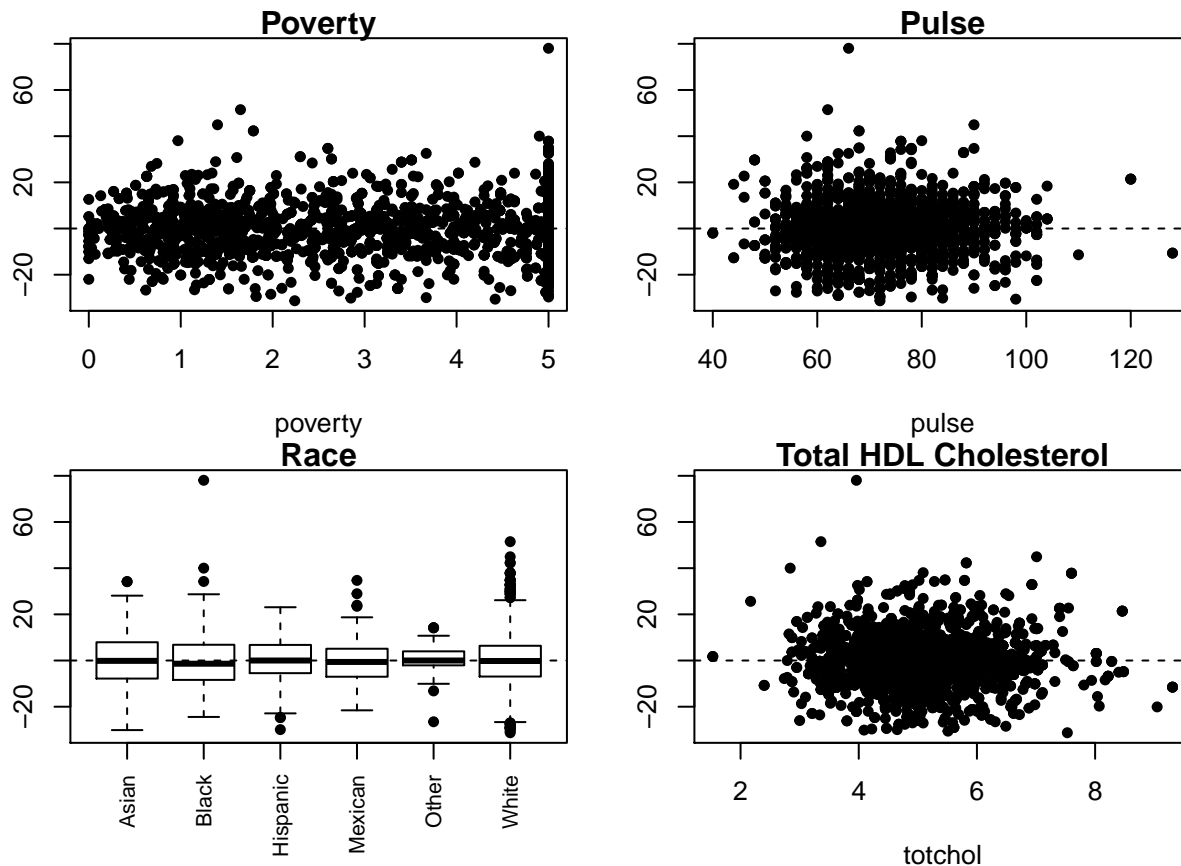
Model adequacy for bpsys

```
##      Rsq Adj.Rsq      Cp      AIC      BIC
## 1 0.2633  0.2494 568.0815 17011.02 17242.74
## 2 0.2621  0.2489 568.3441 17010.48 17230.89
## 3 0.2602  0.2477 570.9761 17011.85 17220.96
## 4 0.4536  0.3672 412.6086 16878.39 18511.71
## 5 0.4493  0.3668 401.8348 16868.99 18428.83
## 6 0.4864  0.4023 315.0000 16765.92 18450.09
```

Fit 6 has lowest AIC and highest R_{adj}^2 , thus it is most adequate.

Residual plots for bpsys





Residual plots seem reasonable with constant variance. Factor predictors have less consistent residuals but they seem to even out. Drinks/day has higher variance for lower values but it seems to be caused by abundance of those who do not drink in a day.

Systolic blood pressure is slightly higher for older people, those who drink more, and those with higher BMI, highest for high school grads (59.051), highest for other races and lowest for Hispanics (incl. Mexicans) and whites, lower for less poor people (-1.3378). Lower and middle classes seems to have lowest bpsys, similar to bpdia. Those who live with their partner or are married have higher bpsys. Pulse has little effect on both blood pressures (-.273 and -.317 respectively).