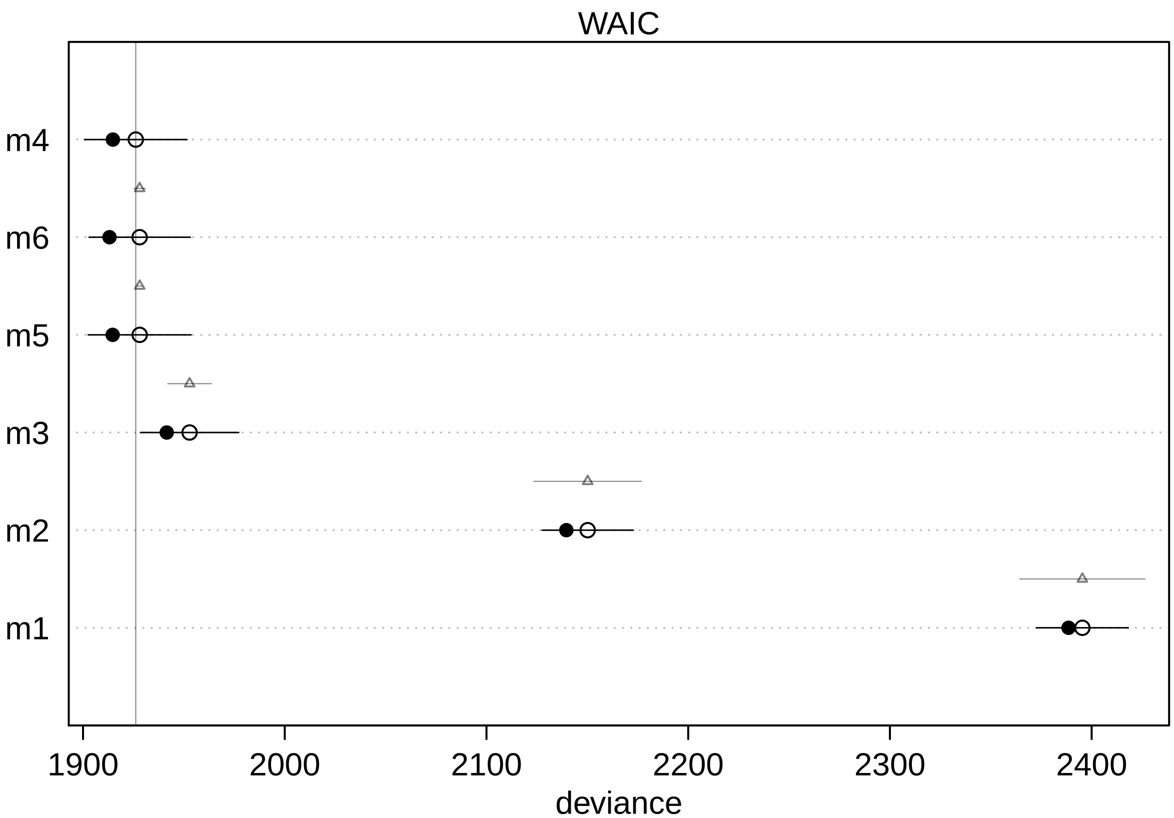
HW5

Matt Kaye

10/14/2018

6H1



	WAIC	pWAIC	dWAIC	weight	SE	dSE
m4	1926.142	5.673	0.000	0.567	25.465	NA
m6	1928.058	7.467	1.916	0.218	25.089	2.881
m5	1928.085	6.699	1.943	0.215	25.534	1.130
m3	1952.826	5.660	26.684	0.000	24.374	10.881
m2	2150.174	5.288	224.032	0.000	22.613	26.735
m1	2395.393	3.418	469.251	0.000	22.808	31.016

According to WAIC, the fourth-, fifth-, and sixth- degree polynomials all perform equally well, and all three perform significantly better than the first-, second-, and third- degree polynomial models. The weights of the diffferent models are roughly the probabilities that each model is the “correct” one.

6H4

```r
##check out this function!!
dev.solver <- function(x, fit) {
  num.terms <- length(coef(fit))-1
  age.powers <- data.frame(matrix(NA, nrow = length(x), ncol = num.terms-1))
  for(i in 1:ncol(age.powers)) {
    age.powers[,i] <- x**i
  }

  mu.list <- data.frame(matrix(NA, nrow = length(x), ncol = num.terms))
  mu.list[,1] <- unname(coef(fit)[1])
  for(i in 2:ncol(mu.list)){
    temp <- age.powers[,i-1]*unname(coef(fit)[i])
    mu.list[,i] <- temp
  }
  sums <- rowSums(mu.list)
  dev <- -2*sum(dnorm(d2$height, mean=sums, sd=coef(fit)[length(coef(fit))], log = TRUE))
  return(dev)
}

fits <- c(m1,m2,m3,m4,m5,m6)
deviances <- NULL

for(fit in fits) {
  deviances <- append(deviances, dev.solver(d2$age, fit))
}

#couldn't figure out how to assign names without brute force
mod.names <- c("m1", "m2", "m3", "m4", "m5", "m6")
deviances <- data.frame(mod.names, deviances)
knitr::kable(deviances, format = "markdown", digits = 3, align = 'c')
```

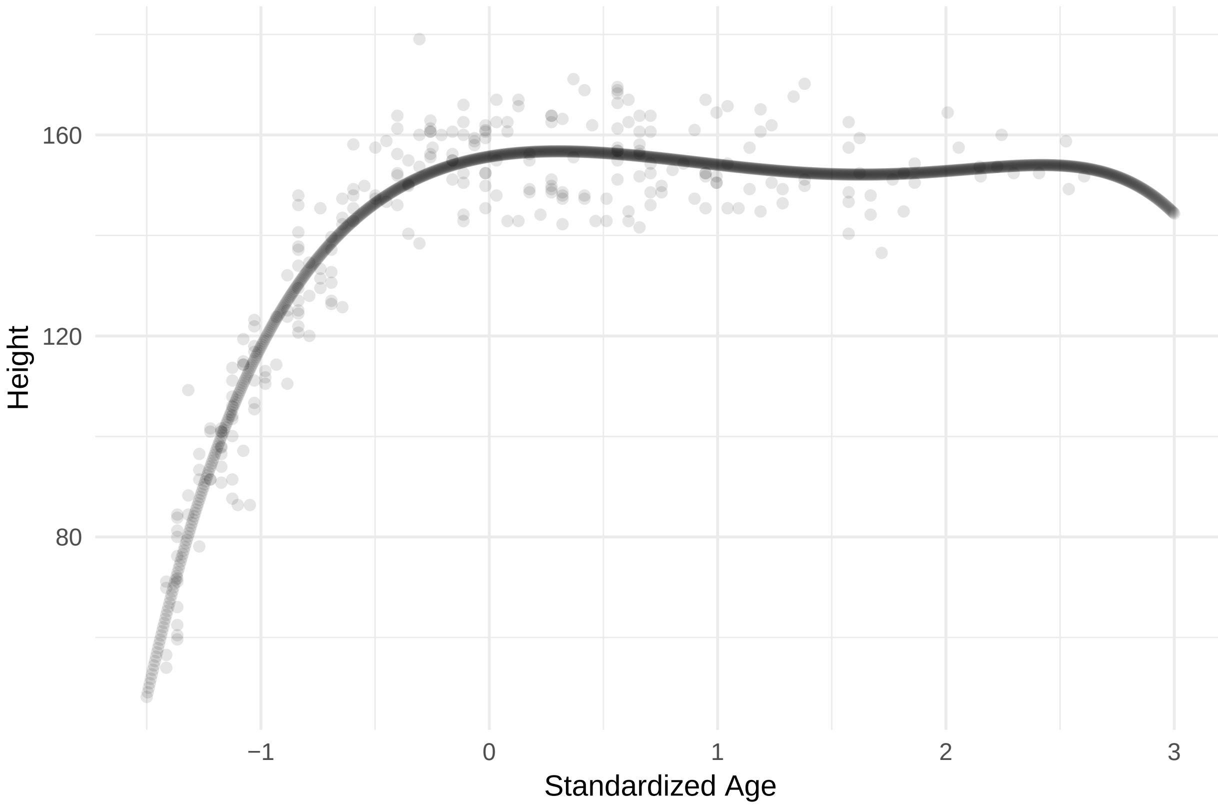
mod.names	deviances
m1	2422.091
m2	2137.512
m3	1932.257
m4	1876.568
m5	1878.390
m6	1877.629

6H5

mod.names	deviances	dDev	WAIC	dWAIC
m1	2422.091	545.523	2395.393	469.251
m2	2137.512	260.944	2150.174	224.032
m3	1932.257	55.689	1952.826	26.684
m4	1876.568	0.000	1926.142	0.000
m5	1878.390	1.822	1928.085	1.943
m6	1877.629	1.061	1928.058	1.916

As before, the fourth-degree polynomial model still makes the best out-of-sample predictions in this case. Additionally, seeing the table above, it is reassuring that the WAIC values are similar to the out-of-sample deviances, which means that WAIC is does a good job of estimating test deviance in this case.

6H6



Points are age, height pairs from d1. polynomial is predicted values of d2 heights given gridded ages

	Mean	StdDev	5.5%	94.5%
a	155.727	0.881	154.319	157.135
b1	7.647	1.659	4.995	10.298
b2	-16.411	2.138	-19.827	-12.994
b3	9.211	2.404	5.368	13.054
b4	-3.116	1.151	-4.955	-1.276
b5	1.259	0.933	-0.233	2.750
b6	-0.260	0.301	-0.740	0.221
sigma	8.243	0.359	7.670	8.817

	Mean	StdDev	5.5%	94.5%
a	156.687	0.790	155.424	157.949
b1	5.963	1.397	3.731	8.196
b2	-19.277	1.151	-21.117	-17.437
b3	12.346	0.911	10.890	13.801
b4	-2.309	0.429	-2.994	-1.624
sigma	8.165	0.350	7.606	8.725

Both the WAIC values and the deviances are almost identical between the model from this question with regularizing priors and the best WAIC model from the first question (the fourth-degree polynomial model). The difference in deviances is -1.473, which is essentially zero. This means that these two models have about equally good out-of-sample performance.