Hw7 Stat 139

Callin Switzer

October 28, 2014

#2 (#26a)

Below are the tables for the linear regression for each country. I've also calculatedd the $\hat{\sigma}$ (sigmaHat) for each country and printed it under its respective table. The values are all the same as Diaplay 7.17 in the book

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5987232938	0.0408047207	14.6728928254	0.0000000000
births$Year	-0.0000428854	0.0000206916	-2.0725982922	0.0442382845

[1] "sigmaHat = 0.0018 for Denmark"

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.6723983751	0.0279195810	24.0833977682	0.0000000000
births$Year	-0.0000808432	0.0000141577	-5.7101958901	0.0000009637

[1] "sigmaHat = 0.00123 for Netherlands"

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.7337857143	0.0548006828	13.3900834282	0.0000000000
births$Year	-0.0001111688	0.0000276770	-4.0166527843	0.0007375947

[1] "sigmaHat = 0.00077 for Canada"

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.6200857143	0.0185987668	33.3401521034	0.0000000000
births$Year	-0.0000542857	0.0000093933	-5.7792120330	0.0000143911

[1] "sigmaHat = 0.00026 for USA"

26b

The t-statistics are shown in the tables above. There is evidence that the proportion of male birth is truly declining. The p-values are highly significant for Netherlands, Candada, and USA, and somewhat significant for Denmark.

26c

The t-statistic is the ratio of the coefficient estimates to their standard errors. The US can have an extreme t-statistic because the standard error is very small, relative to the estimate. This makes the t-statistic more extreme.

26d

The standard error for the estimated slope is smaller in the US than Canada because the points fall closer to the line. Standard error of β_1 depends on $\hat{\sigma}$, which becomes larger when the sum of squared residuals gets larger. The SE(β_1) also depends on the sample variance of the x-values, but those should be the same, since the US and Canada were sampled during the same years.

Code

```
options(xtable.comment = FALSE)
births <- read.csv("data/ex0724.csv")

countries <- colnames(births)[2:5]

mods <- lapply(countries, function(x) summary(lm(births[[x]] ~ births$Year)))

names(mods) <- countries


require(xtable)
x <- "Denmark"

foo <- lm(births[[x]]~births$Year)
sigmaHat <- sqrt(sum((foo$residuals)^2)/length(foo$residuals))
#summary(foo)
#print(xtable(foo))

options("scipen"=10, "digits"=4)
for(i in 1:4){
     sigmaHat <-  sqrt(sum((mods[[i]]$residuals)^2)/(length(mods[[i]]$residuals) - 2))
     sigmaHat <- round(sigmaHat, digits = 5)
     capt = (paste("sigmaHat =", sigmaHat, "for", countries[i]))
     print(xtable(mods[[i]], digits = c(10, 10, 10, 10, 10),
                  caption = capt), floating = F)
     print(paste(capt))
}
```