

Practice02

DSC 384-01a

2024-11-28

```
library(alr4)
```

```
## Loading required package: car
## Loading required package: carData
## Loading required package: effects
## lattice theme set by effectsTheme()
## See ?effectsTheme for details.
```

```
summary(Heights)
```

```
##      mheight      dheight
##  Min.   :55.40   Min.   :55.10
##  1st Qu.:60.80   1st Qu.:62.00
##  Median :62.40   Median :63.60
##  Mean   :62.45   Mean   :63.75
##  3rd Qu.:63.90   3rd Qu.:65.60
##  Max.   :70.80   Max.   :73.10
```

```
library("rstan")
```

```
## Loading required package: StanHeaders
##
## rstan version 2.32.3 (Stan version 2.26.1)
## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
## For within-chain threading using `reduce_sum()` or `map_rect()` Stan functions,
## change `threads_per_chain` option:
## rstan_options(threads_per_chain = 1)
```

```
library("arm")
```

```
## Loading required package: MASS
## Loading required package: Matrix
## Loading required package: lme4
##
## arm (Version 1.14-4, built: 2024-4-1)
## Working directory is /Users/giri/Documents/UTAustin/DSC384/Week12
```

```

##
## Attaching package: 'arm'
## The following object is masked from 'package:rstan':
##
##     traceplot
## The following object is masked from 'package:car':
##
##     logit
library("rstanarm")

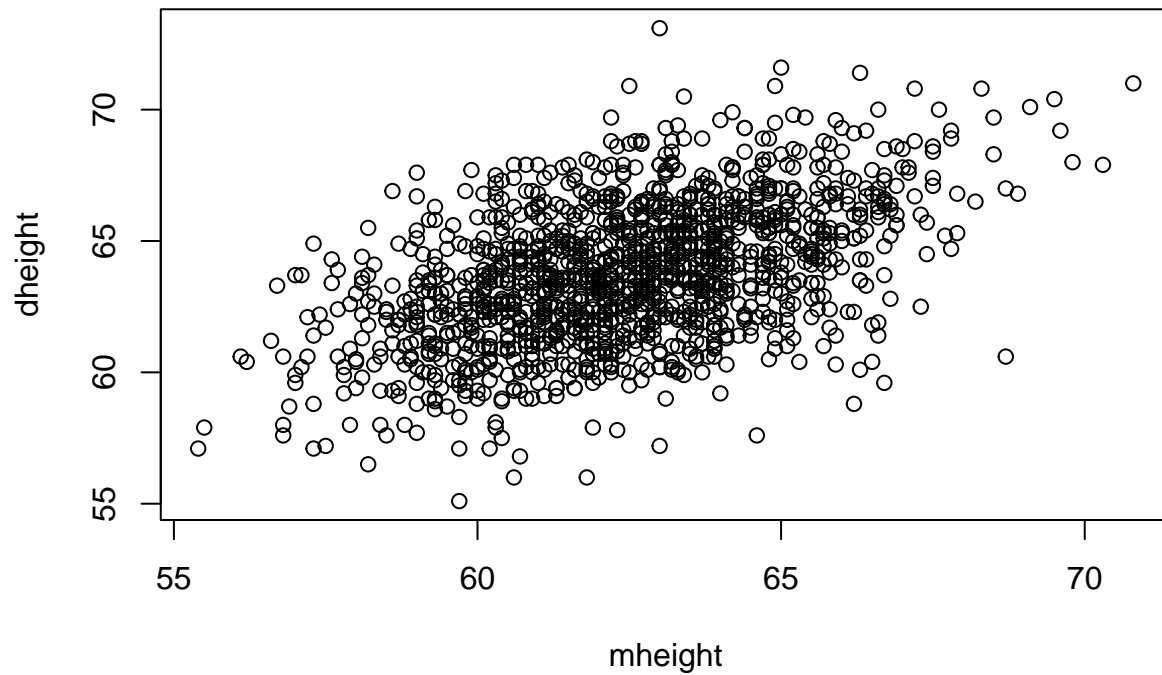
## Loading required package: Rcpp
## This is rstanarm version 2.26.1
## - See https://mc-stan.org/rstanarm/articles/priors for changes to default priors!
## - Default priors may change, so it's safest to specify priors, even if equivalent to the defaults.
## - For execution on a local, multicore CPU with excess RAM we recommend calling
##   options(mc.cores = parallel::detectCores())
##
## Attaching package: 'rstanarm'
## The following objects are masked from 'package:arm':
##
##     invlogit, logit
## The following object is masked from 'package:rstan':
##
##     loo
## The following object is masked from 'package:car':
##
##     logit
library("survey")

## Loading required package: grid
## Loading required package: survival
##
## Attaching package: 'survey'
## The following object is masked from 'package:graphics':
##
##     dotchart
head(Heights)

##      mheight dheight
## 1      59.7      55.1
## 2      58.2      56.5
## 3      60.6      56.0
## 4      60.7      56.8
## 5      61.8      56.0
## 6      55.5      57.9

```

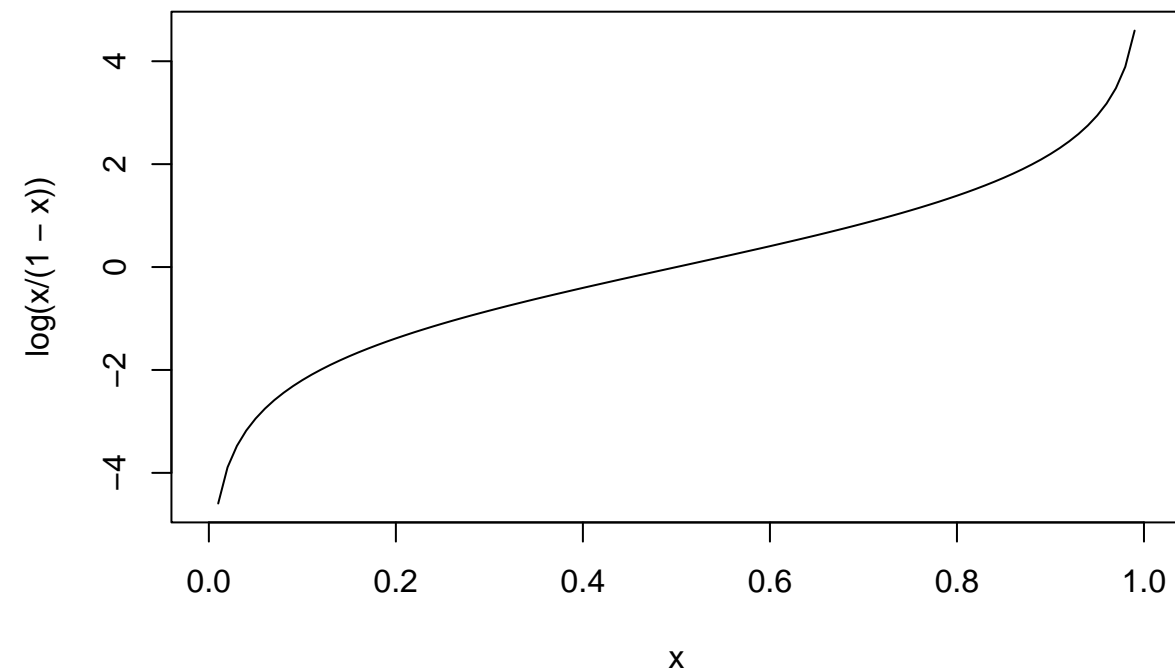
```
plot(dheight ~ mheight , data= Heights)
```



```
dim(Heights)
```

```
## [1] 1375 2
```

```
curve(log(x/(1-x)), from =0.0, to=1.0)
```



```
cc2 <- read.csv("cc2.csv")
head(cc2)
```

```
## momage momrace b.marr momed work.dur prenatal cig booze sex first bw
```

```

## 1      33      3      1      4      1      1 0      0 1      0 1559
## 2      22      1      0      1      0      1 0      1 1      0 2240
## 3      13      1      0      1      0      1 0      0 1      1 1900
## 4      25      1      1      4      1      1 0      0 1      1 1550
## 5      19      1      0      1      0      1 1      0 1      0 2270
## 6      19      1      0      2      1      1 1      1 0      1 1550
##      preterm      age dayskidh ppvtr.36 sample b.state st5 st9 st12 st25 st36 st42
## 1      10 60.79671      31      111      I      5 1 0 0 0 0 0
## 2      3 59.77823      4      81      I      5 1 0 0 0 0 0
## 3      6 59.51540      9      92      I      5 1 0 0 0 0 0
## 4      8 59.18686      50     103      I      5 1 0 0 0 0 0
## 5      5 58.79261      4      81      I      5 1 0 0 0 0 0
## 6      4 58.49692      13     94      I      5 1 0 0 0 0 0
##      st48 st53 income      inneeds unemp.rt st99 treat treat0 b.8state black
## 1      0 0 42500 2.0083910      2 FALSE      1      1      5 0
## 2      0 0 5000 0.3665904      4 FALSE      1      1      5 1
## 3      0 0 12500 0.2660848      5 FALSE      1      1      5 1
## 4      0 0 42500 4.8643700      6 FALSE      1      1      5 1
## 5      0 0 5000 0.4463090      7 FALSE      1      1      5 1
## 6      0 0 12500 0.4463090      5 FALSE      1      1      5 1
##      hispanic white bs.by.aa lths hs ltcoll college v.yng young y.adult late20
## 1      0 1 105 0 0 0 1 0 0 0 0
## 2      0 0 205 1 0 0 0 0 0 0 1 0
## 3      0 0 205 1 0 0 0 1 0 0 0 0
## 4      0 0 205 0 0 0 1 0 0 0 0 1
## 5      0 0 205 1 0 0 0 0 1 0 0 0
## 6      0 0 205 0 1 0 0 0 1 0 0 0
##      older ptc cat ptc0 ptc1 ptc2 ptc3 bwg ethnic educ educ3 state state2
## 1      1 3 0 0 0 1 0 3 4 3 1 0
## 2      0 2 0 0 1 0 1 2 1 1 1 0
## 3      0 3 0 0 0 1 0 2 1 1 1 0
## 4      0 3 0 0 0 1 0 2 4 3 1 0
## 5      0 3 0 0 0 1 1 2 1 1 1 0
## 6      0 2 0 0 1 0 0 2 2 2 1 0
##      state3 neg.bw no.prenatal b.unmarr bwT dayskidT pretermT momageT
## 1      1 941 0 0 3481 3.465736 324 1089
## 2      1 260 0 1 547600 1.609438 121 484
## 3      1 600 0 1 160000 2.302585 196 169
## 4      1 950 0 0 2500 3.931826 256 625
## 5      1 230 0 1 592900 1.609438 169 361
## 6      1 950 0 1 2500 2.639057 144 361

```

```

#a1 <- mean(cc2[cc2$treat == 0,]$bw)
#a2 <- mean(cc2[cc2$treat == 1,]$bw)
#sd1 <- sd(cc2$bw)
#sd2 <- sd(cc2[cc2$treat == 1,]$bw)
#((a1-a2)/sd2)

```

```

a1 <- mean(cc2[cc2$treat == 0,]$dayskidh)
a2 <- mean(cc2[cc2$treat == 1,]$dayskidh)
sd1 <- sd(cc2$dayskidh)
sd2 <- sd(cc2[cc2$treat == 1,]$dayskidh)
((a1-a2)/sd2)

```

```
## [1] -0.9321339
```

```
#my_data[gender == "male" & age <= 20]  
#colnames(cc2)
```

```
colnames(cc2)
```

```
## [1] "momage"      "momrace"      "b.marr"      "momed"      "work.dur"  
## [6] "prenatal"    "cig"          "booze"       "sex"        "first"  
## [11] "bw"          "preterm"      "age"         "dayskidh"   "ppvtr.36"  
## [16] "sample"      "b.state"      "st5"         "st9"         "st12"  
## [21] "st25"        "st36"         "st42"        "st48"        "st53"  
## [26] "income"      "ineeds"       "unemp.rt"    "st99"        "treat"  
## [31] "treat0"      "b.8state"     "black"       "hispanic"    "white"  
## [36] "bs.by.aa"    "lths"         "hs"          "ltcoll"      "college"  
## [41] "v.yng"       "young"        "y.adult"     "late20"      "older"  
## [46] "ptcat"       "ptcat0"       "ptcat1"      "ptcat2"      "ptcat3"  
## [51] "bwg"         "ethnic"       "educ"        "educ3"       "state"  
## [56] "state2"      "state3"       "neg.bw"      "no.prenatal" "b.unmarr"  
## [61] "bwT"         "dayskidT"    "pretermT"    "momageT"
```

```
bs1 <- read.csv("bart_sim.csv")  
head(bs1)
```

```
##   X      X1      X2 Z      Y      p      p_pred  
## 1 1  0.7835386 0.5683032 1 2.766009 0.6178991 0.9110598  
## 2 2 -1.7651494 0.9520432 0 2.985286 0.7758282 0.8476326  
## 3 3 -2.1837289 1.2442086 1 4.014060 0.7983417 0.8525180  
## 4 4  3.3942967 0.7995387 1 1.788870 0.7340758 0.8734638  
## 5 5 -3.6177691 1.3784607 1 4.949849 0.8406391 0.7588949  
## 6 6  0.4991688 0.7305654 0 2.960761 0.6768442 0.9001571
```

```
a1 <- mean(bs1[bs1$Z == 0,]$X1)  
a2 <- mean(bs1[bs1$Z == 1,]$X1)  
sd1 <- sd(bs1$X1)  
sd2 <- sd(bs1[bs1$Z == 1,]$X1)  
(abs(a2-a1)/sd2)
```

```
## [1] 0.07628812
```

```
ps_fit_1 <- stan_glm(Z ~ X1 + X2 + X1:X2, family=binomial(link="logit"), data=bs1, algorithm="optimizing")
```

```
pscores <- apply(posterior_linpred(ps_fit_1), 2, mean)
```

```
bs2 <- bs1  
bs2$pscores <- pscores  
head(bs2)
```

```
##   X      X1      X2 Z      Y      p      p_pred pscores  
## 1 1  0.7835386 0.5683032 1 2.766009 0.6178991 0.9110598 1.994554  
## 2 2 -1.7651494 0.9520432 0 2.985286 0.7758282 0.8476326 1.839665  
## 3 3 -2.1837289 1.2442086 1 4.014060 0.7983417 0.8525180 1.657492  
## 4 4  3.3942967 0.7995387 1 1.788870 0.7340758 0.8734638 1.834098  
## 5 5 -3.6177691 1.3784607 1 4.949849 0.8406391 0.7588949 1.553727  
## 6 6  0.4991688 0.7305654 0 2.960761 0.6768442 0.9001571 1.923865
```

```
(min(bs2[bs2$Z==0,]$pscores))
```

```
## [1] 0.8724252
```

```

(max(bs2[bs2$Z==0,]$pscores))

## [1] 2.504346
(min(bs2[bs2$Z==1,]$pscores))

## [1] 0.3638897
(max(bs2[bs2$Z==1,]$pscores))

## [1] 3.915808
bs3 <- bs2[bs2$pscores >= 0.8329291 & bs2$pscores <= 2.486686,]

a1 <- mean(bs3[bs3$Z == 0,]$X2)
a2 <- mean(bs3[bs3$Z == 1,]$X2)
sd1 <- sd(bs3$X2)
sd2 <- sd(bs3[bs3$Z == 1,]$X2)
(abs(a2-a1)/sd2)

## [1] 0.1133605
(min(bs3$pscores))

## [1] 0.8724252
(max(bs3$pscores))

## [1] 2.480651
(min(bs3$pscores))

## [1] 0.8724252
(max(bs3$pscores))

## [1] 2.480651

```