ASDA3 Analysis Examples Replication Chapter 9

R Code

```
# Data production prior to analysis examples

setwd("P:/ASDA3/Replication R/Chapter 9")
getwd()

# Data production
library (sas7bdat)
library (survey)
library (brms)
library (csSampling)


# NHANES
# library(haven)
# example of how to read in from SAS, not used here
# data <- read_sas("<path to your SAS file>")
# nhanesdata <- read_sas("P:\\ASDA3\\Replication SAS\\nhanes1112_f.sas7bdat")

# load nhanes1112.rdata
load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/nhanes1112.rdata")
nhanesdata <- nhanes1112
summary(nhanesdata)

#create factor variables
nhanesdata$racec <- factor(nhanesdata$ridreth1, levels = 1: 5 , labels =c("Mexican", "Other Hispanic", "White",
"Black", "Other"))
nhanesdata$marcatc <- factor(nhanesdata$marcat, levels = 1: 3, labels =c("Married", "Previously Married", "Never
Married"))
nhanesdata$edcatc <- factor(nhanesdata$edcat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
nhanesdata$bp_catc <- factor(nhanesdata$bp_cat, levels = 1: 4, labels =c("Normal", "Pre-HBP", "Stage 1
HBP","Stage 2 HBP"))
#nhanesdata$agecsq <- (nhanesdata$age * nhanesdata$age)

names(nhanesdata)

nhanessvy2 <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesdata, nest=T)
subnhanes <- subset(nhanessvy2 , age >= 18)
names (nhanessvy2)


# NCSR-R
load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/ncsr.rdata")
names(ncsr)
summary(ncsr)

#create factor versions with labels
ncsr$racec <- factor(ncsr$racecat, levels = 1: 4, labels =c("Other", "Hispanic", "Black", "White"))
ncsr$mar3catc <- factor(ncsr$mar3cat, levels = 1: 3, labels =c("Married", "Previously Married", "Never
Married"))
ncsr$ed4catc <- factor(ncsr$ed4cat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
ncsr$ag4catc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
ncsr$mdec <- factor(ncsr$mde, level = 1:2, labels=c("No","Yes"))

# part 1 data n=9282
ncsrsvyp1 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtsh, data=ncsr, nest=T)
```

```
names (ncsrsvyp1)

# part 2 data n=5692
ncsrp2 <- subset(ncsr, !is.na(ncsrwtlg))
ncsrsvyp2 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtlg, data=ncsrp2, nest=T)
names (ncsrsvyp2)

# rescale weight to sum to population in 2001
ncsr$popweight <- (ncsr$ncsrwtsh*(209128094/9282))

ncsrsvypop <- svydesign(strata=~sestrat, id=~seclustr, weights=~popweight, data=ncsr, nest=T)
summary(ncsrsvypop)


# HRS
# HH and respondent weights are needed plus financial respondent indicator for hh level analysis
load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/hrs12.rdata")
names(hrs12)

hrssvyhh <- svydesign(strata=~stratum, id=~secu, weights=~nwgthh , data=hrs12, nest=T)
summary(hrssvyhh)

hrssvysub <-subset(hrssvyhh, nfinr==1)
summary(hrssvysub)

hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
summary(hrssvyr)


# Chapter 9 Analysis Examples

# Figure 9.2 Bar chart of work status NCS-R data
fig92 <- svymean( ~factor(wkstat3c), ncsrsvyp2, na.rm=T)
fig92

png("P:/ASDA3/Replication R/Chapter 9/fig92_bar.png")
barplot(fig92, legend=c("Employed", "Unemployed", "NLF") , col=c("black", "grey60", "blue"))
dev.off()

# Tests for Potential Predictors of Work Status
svychisq(~wkstat3c +sex, ncsrsvyp2, statistic="F")
svychisq(~wkstat3c +ald, ncsrsvyp2, statistic="F")
svychisq(~wkstat3c +mde, ncsrsvyp2, statistic="F")
svychisq(~wkstat3c +ed4cat, ncsrsvyp2, statistic="F")
svychisq(~wkstat3c +ag4cat, ncsrsvyp2, statistic="F")
svychisq(~wkstat3c +mar3cat, ncsrsvyp2, statistic="F")

# Note: Use svyVGAM Package for the Multinomial model, install/load package
library(svyVGAM)

# run model for Table 9.2
ex9_2_6 <- svy_vglm(wkstat3c ~ factor(ag4cat) + factor(sex) + ald + mde + factor(ed4cat) +
factor(mar3cat), family=multinomial(refLevel=1), design=ncsrsvyp2)
summary(ex9_2_6)
coef(ex9_2_6)
confint(ex9_2_6)
exp(coef(ex9_2_6))

# Note that intercepts and signs differ from other software due to differences in how reference
```

```
# groups/intercepts are handled but the conclusions are similar

# NOTE: MARGINAL EFFECTS EXAMPLE ARE FORTHCOMING, CHECK WEBSITE FOR UPDATES!

# Alternate Bayesian approach for Section 9.2.6.1
library(csSampling)

# Part 2 of the NCS-R survey, use correct weight ncsrwtlg in filter and design set up
ncsr$marcatc <- factor(ncsr$mar3cat, levels = 1:3, labels =c("Married", "Previously Married", "Never Married"))
ncsr$edcatc <-factor(ncsr$ed4cat, levels = 1:4, labels =c("0-11", "12", "13-15","16+"))
ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
ncsr$sexc <- relevel(ncsr$sexc, ref="Female")
ncsr$agcatc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
ncsr$wkstatc <- factor(ncsr$wkstat3c, levels = 1:3, labels=c("Emp", "Unemp", "NLF"))

ncsr.lg <- ncsr[!is.na(ncsr$ncsrwtlg),]

# Create complete data set with variables of interest
ncsr.red <- ncsr.lg[, c("wkstatc", "marcatc", "edcatc", "sexc", "agcatc", "ald", "mde", "seclustr", "sestrat",
"ncsrwtlg")]
ncsr.red <- ncsr.red[complete.cases(ncsr.red),]

# Normalize weights to match what is done for Stan modeling
ncsr.red$wtsc <- ncsr.red$ncsrwtlg / mean(ncsr.red$ncsrwtlg)

# Survey design object
ncsr.des <- svydesign(id = ~seclustr, strata = ~sestrat, weights = ~wtsc, nest = T, data = ncsr.red)

# Bayesian approach, flat prior
set.seed(41279)
model_formula <- formula("wkstatc|weights(wtsc) ~ sexc + ald + mde + edcatc + agcatc + marcatc")
mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), data = ncsr.red, family =
categorical(link = "logit"))
mod.brms$stan_fit


# Ordinal Regression with Russian Federation Data, Section 9.3.6
load ("P:\\ASDA3\\Data Sets for Analysis Examples and Stata R Code\\ess6_russia.rdata")

# Prepare needed outcome variable
ess6_russia$stflife2[ess6_russia$stflife >= 0 & ess6_russia$stflife <= 1] <- 1
ess6_russia$stflife2[ess6_russia$stflife >= 2 & ess6_russia$stflife <= 4] <- 2
ess6_russia$stflife2[ess6_russia$stflife == 5] <- 3
ess6_russia$stflife2[ess6_russia$stflife >= 6 & ess6_russia$stflife <= 8] <- 4
ess6_russia$stflife2[ess6_russia$stflife >= 9 & ess6_russia$stflife <= 10] <- 5

rfdata_c9 <- ess6_russia
summary(rfdata_c9)

# Create factor variable
rfdata_c9$marcatc <- factor(rfdata_c9$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))
str(rfdata_c9)

rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~pspwght, data=rfdata_c9, nest=T)

ex936 <- svymean(~factor(stflife2), design=rfsvy, na.rm=T, se=T, ci=T, keep.vars=T)
(ex936)

png("P:/ASDA3/Replication R/Chapter 9/fig96_bar2.png")
```

```
barplot(ex936, legend=c("0-1", "2-4", "5", "6-8", "9-10") ,
 col=c("black", "grey60", "blue", "red", "green"))
dev.off()

# Ordinal logistic using satisfaction with life, Tables 9.6 and 9.7
summary(ex936_ordinal <- svyolr (factor(stflife2) ~ factor(agecat) + male + marcatc, design=rfsvy))
#exp for odds ratio
exp(ex936_ordinal$coef)
# Note test of Proptional Odds not included but there are ways to extend svyolr, see T. Lumley notes online for
details

# Bayesian approach Example 9.3.6
# create complete data set with variables of interest

ess.red <- rfdata_c9[, c("stflife2", "agecat", "marcat", "male", "psu", "stratify", "pspwght")]
ess.red <- ess.red[complete.cases(ess.red),]

# need to normalize NCS-R weights to match what is done for Stan modeling
ess.red$wtsc <- ess.red$pspwght / mean(ess.red$pspwght)

# survey design object
ess.des <- svydesign(id = ~psu, strata = ~stratify, weights = ~wtsc, nest = T, data = ess.red)

# Bayesian approach, flat prior
set.seed(41279)
model_formula <- formula("stflife2|weights(wtsc) ~ factor(agecat) + factor(marcat) + male")

# Run model, note the family statement needs the below code, differs from the code in the book (per Matt
Williams)
mod.brms <- cs_sampling_brms(svydes = ess.des, brmsmod = brmsformula(model_formula),
 data = ess.red, family = brms::cumulative(link="logit"))
mod.brms$stan_fit

# Count Outcomes Section 9.4
# HRS data for Poisson Regression using Number of Falls

# Histogram of number of falls past 24 months
png("P:/ASDA3/Replication R/Chapter 9/fig96_bar.png")
svyhist(~numfalls24 , subset (hrssvyr, nage >=65), main="", col="grey80", xlab ="Histogram of Number of Falls
Past 24 Months")
dev.off()

# Poisson model
summary(hrs12)

# Create variables for example
hrs12$nage_c=(hrs12$nage - 74.5)
hrs12$bmi_c=(hrs12$r11bmi - 27.7)
hrs12$male=(hrs12$gender ==1)

# Create Design Object and then subset for 65 Plus
hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
summary(hrssvyr)
hrssvy65 <- subset(hrssvyr, age65p==1)

# Table 9.10
summary(ex947_poisson <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c, design=hrssvy65,
family=quasipoisson(log)))
coefficients <- exp(coef(ex947_poisson))
```

```
coefficients
ci_exp <- exp(confint(ex947_poisson))
ci_exp

# Negative Binomial via MASS with svyglm
library(MASS)

# negative binomial model to compare to Poisson Table 9.10
ex947_negb <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c,
   design=hrssvy65, family=negative.binomial(theta = 1))
coef <- exp(coef(ex947_negb))
coef
ci_exp <- exp(confint(ex947_negb))
ci_exp

# Zero Inflated Negative Binomial not easily done but using extensions to R tools,
# can be done using a Frequentist or
# Bayesian Method, check our website for more examples to come for other models
```

R Results

```
> # Chapter 9 ASDA3 Analysis Examples Replication
>
> # Data production prior to analysis examples
> setwd("P:/ASDA3/Replication R/Chapter 9")
> getwd()
[1] "P:/ASDA3/Replication R/Chapter 9"
>
> # Data production
> library (sas7bdat)
> library (survey)
> library (brms)
> library (csSampling)
>
>
> # load nhanes1112.rdata
> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/nhanes1112.rdata")
> nhanesdata <- nhanes1112
> summary(nhanesdata)
      seqn          ridstatr        riagendr        ridreth1        dmdmartl         wtint2yr     
 Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   : 1.000   Min.   :  3321  
 1st Qu.:64600   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 1.000   1st Qu.: 11352  
 Median :67039   Median :2.000   Median :2.000   Median :3.000   Median : 2.000   Median : 18098  
 Mean   :67039   Mean   :1.957   Mean   :1.502   Mean   :3.229   Mean   : 2.749   Mean   : 31426  
 3rd Qu.:69477   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 5.000   3rd Qu.: 34887  
 Max.   :71916   Max.   :2.000   Max.   :2.000   Max.   :5.000   Max.   :99.000   Max.   :220233  
                                                                 NA's   :4196                     
    wtmec2yr         sdmvpsu         sdmvstra          indfmpir          bpxsy1          bpxdi1     
 Min.   :     0   Min.   :1.000   Min.   : 90.00   Min.   :0.000   Min.   : 74.0   Min.   :  0.0  
 1st Qu.: 11174   1st Qu.:1.000   1st Qu.: 92.00   1st Qu.:0.860   1st Qu.:106.0   1st Qu.: 60.0  
 Median : 18090   Median :2.000   Median : 96.00   Median :1.630   Median :116.0   Median : 68.0  
 Mean   : 31426   Mean   :1.643   Mean   : 95.87   Mean   :2.205   Mean   :119.2   Mean   : 66.9  
 3rd Qu.: 34792   3rd Qu.:2.000   3rd Qu.: 99.00   3rd Qu.:3.580   3rd Qu.:128.0   3rd Qu.: 76.0  
 Max.   :222580   Max.   :3.000   Max.   :103.00   Max.   :5.000   Max.   :238.0   Max.   :120.0  
                                                   NA's   :840     NA's   :3000    NA's   :3000   
     bpxsy2          bpxdi2           bpxsy3          bpxdi3           bpxsy4          bpxdi4      
 Min.   : 74.0   Min.   :  0.00   Min.   : 74.0   Min.   :  0.00   Min.   : 78    Min.   :  0.00  
 1st Qu.:106.0   1st Qu.: 58.00   1st Qu.:106.0   1st Qu.: 58.00   1st Qu.:104    1st Qu.: 64.00  
 Median :116.0   Median : 68.00   Median :116.0   Median : 68.00   Median :116    Median : 72.00  
 Mean   :118.7   Mean   : 66.28   Mean   :118.2   Mean   : 65.91   Mean   :119    Mean   : 71.78  
 3rd Qu.:128.0   3rd Qu.: 76.00   3rd Qu.:128.0   3rd Qu.: 76.00   3rd Qu.:130    3rd Qu.: 82.00  
 Max.   :234.0   Max.   :134.00   Max.   :232.0   Max.   :128.00   Max.   :226    Max.   :130.00  
 NA's   :2848    NA's   :2848     NA's   :2839    NA's   :2839     NA's   :9309   NA's   :9309    
     bmxbmi          lbxtc           age18p          irregular          edcat            age      
 Min.   :12.40   Min.   : 59.0   Min.   :0.0000   Min.   :0.0000   Min.   :1.000   Min.   : 0.0  
 1st Qu.:19.30   1st Qu.:154.0   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.000   1st Qu.: 9.0  
 Median :24.50   Median :179.0   Median :1.0000   Median :0.0000   Median :2.000   Median :26.0  
 Mean   :25.34   Mean   :183.2   Mean   :0.6011   Mean   :0.0134   Mean   :2.102   Mean   :31.4  
 3rd Qu.:29.80   3rd Qu.:209.0   3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:3.000   3rd Qu.:52.0  
 Max.   :82.10   Max.   :523.0   Max.   :1.0000   Max.   :1.0000   Max.   :4.000   Max.   :80.0  
 NA's   :1154    NA's   :2768                     NA's   :729      NA's   :1602                  
     marcat          pre_hibp          bp_cat          ag1829          ag3044           ag4559      
 Min.   :1.000   Min.   :0.0000   Min.   :1.000   Min.   :0.000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1.000   1st Qu.:0.0000   1st Qu.:1.000   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.0000  
 Median :1.000   Median :0.0000   Median :1.000   Median :1.000   Median :0.0000   Median :0.0000  
 Mean   :1.652   Mean   :0.4357   Mean   :1.546   Mean   :0.532   Mean   :0.1461   Mean   :0.1384  
 3rd Qu.:2.000   3rd Qu.:1.0000   3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:0.0000   3rd Qu.:0.0000  
 Max.   :3.000   Max.   :1.0000   Max.   :4.000   Max.   :1.000   Max.   :1.0000   Max.   :1.0000  
 NA's   :4203    NA's   :2701     NA's   :2701                                                     
```

```
      ag60             mex             othhis           white            black           other
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.000   Min.   :0.0000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000
 Median :0.0000   Median :0.0000   Median :0.0000   Median :0.0000   Median :0.000   Median :0.0000
 Mean   :0.1836   Mean   :0.1389   Mean   :0.1103   Mean   :0.3047   Mean   :0.275   Mean   :0.1711
 3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:0.0000
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.000   Max.   :1.0000

     lbdhdd          lbdhddsi        lbdtcsi
 Min.   : 14.00   Min.   :0.360   Min.   : 1.530
 1st Qu.: 43.00   1st Qu.:1.110   1st Qu.: 3.980
 Median : 51.00   Median :1.320   Median : 4.630
 Mean   : 52.63   Mean   :1.361   Mean   : 4.738
 3rd Qu.: 60.00   3rd Qu.:1.550   3rd Qu.: 5.400
 Max.   :175.00   Max.   :4.530   Max.   :13.520
 NA's   :2767     NA's   :2767    NA's   :2768
>
> #create factor variables
> nhanesdata$racec <- factor(nhanesdata$ridreth1, levels = 1: 5 , labels =c("Mexican", "Other Hispanic",
"White", "Black", "Other"))
> nhanesdata$marcatc <- factor(nhanesdata$marcat, levels = 1: 3, labels =c("Married", "Previously Married",
"Never Married"))
> nhanesdata$edcatc <- factor(nhanesdata$edcat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
> nhanesdata$bp_catc <- factor(nhanesdata$bp_cat, levels = 1: 4, labels =c("Normal", "Pre-HBP", "Stage 1
HBP","Stage 2 HBP"))
> #nhanesdata$agecsq <- (nhanesdata$age * nhanesdata$age)
>
> names(nhanesdata)
 [1] "seqn"      "ridstatr"  "riagendr"  "ridreth1"  "dmdmartl"  "wtint2yr"  "wtmec2yr"  "sdmvpsu"
 [9] "sdmvstra"  "indfmpir"  "bpxsy1"    "bpxdi1"    "bpxsy2"    "bpxdi2"    "bpxsy3"    "bpxdi3"
[17] "bpxsy4"    "bpxdi4"    "bmxbmi"    "lbxtc"     "age18p"    "irregular" "edcat"     "age"
[25] "marcat"    "pre_hibp"  "bp_cat"    "ag1829"    "ag3044"    "ag4559"    "ag60"      "mex"
[33] "othhis"    "white"     "black"     "other"     "lbdhdd"    "lbdhddsi"  "lbdtcsi"   "racec"
[41] "marcatc"   "edcatc"    "bp_catc"
>
> nhanessvy2 <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesdata, nest=T)
> subnhanes <- subset(nhanessvy2 , age >= 18)
> names (nhanessvy2)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"  "fpc"
[9] "pps"
>
>
>
> # NCSR-R
> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/ncsr.rdata")
> names(ncsr)
 [1] "caseid"   "dsm_so"   "age"      "region"   "mar3cat"  "ed4cat"   "obese6ca" "ncsrwtsh" "ncsrwtlg"
[10] "sex"      "wkstat3c" "sestrat"  "seclustr" "ag4cat"   "racecat"  "mde"      "ald"      "sexf"
[19] "sexm"
>
> #create factor versions with labels
> ncsr$racec <- factor(ncsr$racecat, levels = 1: 4, labels =c("Other", "Hispanic", "Black", "White"))
> ncsr$mar3catc <- factor(ncsr$mar3cat, levels = 1: 3, labels =c("Married", "Previously Married", "Never
Married"))
> ncsr$ed4catc <- factor(ncsr$ed4cat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
> ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
> ncsr$ag4catc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
> ncsr$mdec <- factor(ncsr$mde, level = 1:2, labels=c("No","Yes"))
>
```

```
> # part 1 data n=9282
> ncsrsvyp1 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtsh, data=ncsr, nest=T)
> names (ncsrsvyp1)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"  "fpc"       
[9] "pps"       
>
> # part 2 data n=5692
> ncsrp2 <- subset(ncsr, !is.na(ncsrwtlg))
> ncsrsvyp2 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtlg, data=ncsrp2, nest=T)
> names (ncsrsvyp2)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"  "fpc"       
[9] "pps"       
>
> # rescale weight to sum to population in 2001
> ncsr$popweight <- (ncsr$ncsrwtsh*(209128094/9282))
>
> ncsrsvypop <- svydesign(strata=~sestrat, id=~seclustr, weights=~popweight, data=ncsr, nest=T)
> summary(ncsrsvypop)
Stratified 1 - level Cluster Sampling design (with replacement)
With (84) clusters.
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~popweight, 
    data = ncsr, nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
6.220e-06 3.821e-05 5.089e-05 5.313e-05 6.433e-05 2.621e-04 
Stratum Sizes: 
            1   2   3   4   5   6   7   8   9  10  11  12 13 14  15  16 17 18  19  20  21  22  23  24  25
obs        91 108 123 134 124 113 188 157 140 126 107 108 98 79 109 107 74 86 113 260 309 297 269 286 378
design.PSU  2   2   2   2   2   2   2   2   2   2   2   2  2  2   2   2  2  2   2   2   2   2   2   2   2
actual.PSU  2   2   2   2   2   2   2   2   2   2   2   2  2  2   2   2  2  2   2   2   2   2   2   2   2
            26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42
obs        272 279 214 336 239 393 289 149 361 343 305 333 366 401 313 346 359
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
Data variables:
 [1] "caseid"    "dsm_so"    "age"       "region"    "mar3cat"   "ed4cat"    "obese6ca"  "ncsrwtsh" 
 [9] "ncsrwtlg"  "sex"       "wkstat3c"  "sestrat"   "seclustr"  "ag4cat"    "racecat"   "mde"      
[17] "ald"       "sexf"      "sexm"      "racec"     "mar3catc"  "ed4catc"   "sexc"      "ag4catc"  
[25] "mdec"      "popweight"
>
>
> # HRS
> # HH and respondent weights are needed plus financial respondent indicator for hh level analysis
> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/hrs12.rdata")
> names(hrs12)
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"     "secu"       "stratum"   
 [9] "nwgthh"     "nwgtr"      "h11atota"   "h11itot"    "marcat"     "edcat"      "racecat"    "diabetes"  
[17] "numfalls24" "age65p"     "arthritis" 
>
> hrssvyhh <- svydesign(strata=~stratum, id=~secu, weights=~nwgthh , data=hrs12, nest=T)
> summary(hrssvyhh)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
svydesign(strata = ~stratum, id = ~secu, weights = ~nwgthh, data = hrs12, 
    nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
5.594e-05 1.806e-04 2.978e-04       Inf 5.727e-04       Inf 
Stratum Sizes: 
```

```
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23 24 25
obs        352 627 366 333 276 334 384 403 367 334 261 370 209 159 177 136 276 185 295 371 332 182 107 68 83
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2
            26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49
obs        302 588 432 499 533 514 457 560 483 306 542 196 370 366 669 510 489 777 691 601 687 447 297 333
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            50  51  52  53 54  55  56
obs        685 548 153 145 56 179 152
design.PSU   2   2   2   2  2   2   2
actual.PSU   2   2   2   2  2   2   2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"     "secu"       "stratum"   
 [9] "nwgthh"     "nwgtr"      "h11atota"   "h11itot"    "marcat"     "edcat"      "racecat"    "diabetes"  
[17] "numfalls24" "age65p"     "arthritis" 
>
> hrssvysub <-subset(hrssvyhh, nfinr==1)
> summary(hrssvysub)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
subset(hrssvyhh, nfinr == 1)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
5.594e-05 1.959e-04 3.110e-04       Inf 5.961e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15 16  17  18  19  20  21  22 23 24 25
obs        257 409 247 228 194 245 281 275 253 228 185 253 134 105 125 96 192 123 204 273 233 129 77 52 61
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2   2   2   2   2  2  2  2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2   2   2   2   2  2  2  2
            26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49
obs        217 424 308 350 364 364 314 405 339 217 358 136 258 242 454 345 349 498 445 406 467 308 203 244
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            50  51  52 53 54  55 56
obs        463 387 111 96 40 124 96
design.PSU   2   2   2  2  2   2  2
actual.PSU   2   2   2  2  2   2  2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"     "secu"       "stratum"   
 [9] "nwgthh"     "nwgtr"      "h11atota"   "h11itot"    "marcat"     "edcat"      "racecat"    "diabetes"  
[17] "numfalls24" "age65p"     "arthritis" 
>
> hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
> summary(hrssvyr)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
svydesign(strata = ~stratum, id = ~secu, weights = ~nwgtr, data = hrs12, 
    nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
5.034e-05 1.757e-04 2.874e-04       Inf 5.754e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23 24 25
obs        352 627 366 333 276 334 384 403 367 334 261 370 209 159 177 136 276 185 295 371 332 182 107 68 83
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2
            26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49
obs        302 588 432 499 533 514 457 560 483 306 542 196 370 366 669 510 489 777 691 601 687 447 297 333
```

```
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
              50  51  52  53 54  55  56
obs         685 548 153 145 56 179 152
design.PSU   2   2   2   2  2   2   2
actual.PSU   2   2   2   2  2   2   2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"     "secu"       "stratum"
 [9] "nwgthh"     "nwgtr"      "h11atota"   "h11itot"    "marcat"     "edcat"      "racecat"    "diabetes"
[17] "numfalls24" "age65p"     "arthritis"
```

```
> # Chapter 9 Analysis Examples
> # Figure 9.2 Bar chart of work status NCS-R data
> fig92 <- svymean( ~factor(wkstat3c), ncsrsvyp2, na.rm=T)
> fig92
                      mean     SE
factor(wkstat3c)1 0.647848 0.0103
factor(wkstat3c)2 0.051139 0.0052
factor(wkstat3c)3 0.301013 0.0094
>
> # Tests for Potential Predictors of Work Status
> svychisq(~wkstat3c +sex, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + sex, ncsrsvyp2, statistic = "F")
F = 27.329, ndf = 1.875, ddf = 78.748, p-value = 2.171e-09

> svychisq(~wkstat3c +ald, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + ald, ncsrsvyp2, statistic = "F")
F = 3.1249, ndf = 1.7248, ddf = 72.4410, p-value = 0.05716

> svychisq(~wkstat3c +mde, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + mde, ncsrsvyp2, statistic = "F")
F = 4.6693, ndf = 1.7348, ddf = 72.8614, p-value = 0.01605

> svychisq(~wkstat3c +ed4cat, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + ed4cat, ncsrsvyp2, statistic = "F")
F = 27.64, ndf = 5.1457, ddf = 216.1180, p-value < 2.2e-16

> svychisq(~wkstat3c +ag4cat, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + ag4cat, ncsrsvyp2, statistic = "F")
F = 113.49, ndf = 4.9646, ddf = 208.5128, p-value < 2.2e-16

> svychisq(~wkstat3c +mar3cat, ncsrsvyp2, statistic="F")

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~wkstat3c + mar3cat, ncsrsvyp2, statistic = "F")
F = 23.124, ndf = 3.1985, ddf = 134.3368, p-value = 1.229e-12

> # Note: Use svyVGAM Package for the Multinomial model, install/load package
> library(svyVGAM)
Loading required package: VGAM
Loading required package: stats4
Loading required package: splines

Attaching package: ‘VGAM’
```

```
The following objects are masked from ‘package:brms’:

    acat, cratio, cumulative, dfrechet, dirichlet, exponential, frechet, geometric, lognormal,
    multinomial, negbinomial, pfrechet, qfrechet, rfrechet, s, sratio

The following object is masked from ‘package:survey’:

    calibrate

Warning messages:
1: package ‘svyVGAM’ was built under R version 4.4.3
2: package ‘VGAM’ was built under R version 4.4.3


> # run model for Table 9.2
> ex9_2_6 <- svy_vglm(wkstat3c ~ factor(ag4cat) + factor(sex) + ald + mde + factor(ed4cat) +
+ factor(mar3cat), family=multinomial(refLevel=1), design=ncsrsvyp2)
> summary(ex9_2_6)
svy_vglm.survey.design(wkstat3c ~ factor(ag4cat) + factor(sex) +
    ald + mde + factor(ed4cat) + factor(mar3cat), family = multinomial(refLevel = 1),
    design = ncsrsvyp2)
Stratified 1 - level Cluster Sampling design (with replacement)
With (84) clusters.
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg,
    data = ncsrp2, nest = T)
                        Coef        SE       z         p
(Intercept):1      -2.036998  0.268139 -7.5968 3.035e-14
(Intercept):2      -1.019730  0.156598 -6.5118 7.427e-11
factor(ag4cat)2:1  -0.852391  0.294547 -2.8939 0.0038048
factor(ag4cat)2:2  -0.316445  0.128783 -2.4572 0.0140025
factor(ag4cat)3:1  -0.837701  0.258104 -3.2456 0.0011721
factor(ag4cat)3:2   0.064993  0.170881  0.3803 0.7036929
factor(ag4cat)4:1   1.828395  0.294691  6.2044 5.489e-10
factor(ag4cat)4:2   2.380607  0.173400 13.7290 < 2.2e-16
factor(sex)2:1      1.393197  0.197640  7.0492 1.800e-12
factor(sex)2:2      0.640256  0.110040  5.8184 5.941e-09
ald:1              -0.163781  0.356951 -0.4588 0.6463534
ald:2               0.333248  0.130242  2.5587 0.0105071
mde:1              -0.139756  0.157245 -0.8888 0.3741209
mde:2               0.098522  0.087975  1.1199 0.2627619
factor(ed4cat)2:1  -0.847040  0.235396 -3.5984 0.0003202
factor(ed4cat)2:2  -0.651401  0.141019 -4.6192 3.851e-06
factor(ed4cat)3:1  -1.365302  0.257515 -5.3018 1.146e-07
factor(ed4cat)3:2  -0.916942  0.146494 -6.2593 3.868e-10
factor(ed4cat)4:1  -1.730957  0.310488 -5.5750 2.476e-08
factor(ed4cat)4:2  -1.229501  0.159589 -7.7042 1.317e-14
factor(mar3cat)2:1 -0.589903  0.225228 -2.6191 0.0088152
factor(mar3cat)2:2 -0.052263  0.105011 -0.4977 0.6187010
factor(mar3cat)3:1 -2.784566  0.380263 -7.3227 2.430e-13
factor(mar3cat)3:2  0.552789  0.132363  4.1763 2.963e-05
> coef(ex9_2_6)
     (Intercept):1      (Intercept):2  factor(ag4cat)2:1  factor(ag4cat)2:2  factor(ag4cat)3:1
       -2.03699802        -1.01972961        -0.85239068        -0.31644499        -0.83770055
 factor(ag4cat)3:2  factor(ag4cat)4:1  factor(ag4cat)4:2     factor(sex)2:1     factor(sex)2:2
        0.06499308         1.82839494         2.38060666         1.39319687         0.64025553
             ald:1              ald:2              mde:1              mde:2  factor(ed4cat)2:1
       -0.16378133         0.33324772        -0.13975600         0.09852204        -0.84703977
 factor(ed4cat)2:2  factor(ed4cat)3:1  factor(ed4cat)3:2  factor(ed4cat)4:1  factor(ed4cat)4:2
```

```
       -0.65140124        -1.36530168        -0.91694203        -1.73095701        -1.22950065
factor(mar3cat)2:1 factor(mar3cat)2:2 factor(mar3cat)3:1 factor(mar3cat)3:2
       -0.58990261        -0.05226339        -2.78456611         0.55278909
> confint(ex9_2_6)
                         2.5 %      97.5 %
(Intercept):1      -2.56254040 -1.51145565
(Intercept):2      -1.32665625 -0.71280297
factor(ag4cat)2:1  -1.42969183 -0.27508953
factor(ag4cat)2:2  -0.56885472 -0.06403526
factor(ag4cat)3:1  -1.34357549 -0.33182562
factor(ag4cat)3:2  -0.26992840  0.39991455
factor(ag4cat)4:1   1.25081059  2.40597929
factor(ag4cat)4:2   2.04074982  2.72046351
factor(sex)2:1      1.00583012  1.78056362
factor(sex)2:2      0.42458200  0.85592907
ald:1              -0.86339287  0.53583022
ald:2               0.07797766  0.58851779
mde:1              -0.44794980  0.16843781
mde:2              -0.07390577  0.27094984
factor(ed4cat)2:1  -1.30840783 -0.38567172
factor(ed4cat)2:2  -0.92779316 -0.37500933
factor(ed4cat)3:1  -1.87002102 -0.86058234
factor(ed4cat)3:2  -1.20406408 -0.62981998
factor(ed4cat)4:1  -2.33950148 -1.12241254
factor(ed4cat)4:2  -1.54228852 -0.91671278
factor(mar3cat)2:1 -1.03134068 -0.14846455
factor(mar3cat)2:2 -0.25808212  0.15355533
factor(mar3cat)3:1 -3.52986728 -2.03926494
factor(mar3cat)3:2  0.29336273  0.81221545
> exp(coef(ex9_2_6))
     (Intercept):1      (Intercept):2  factor(ag4cat)2:1  factor(ag4cat)2:2  factor(ag4cat)3:1
        0.13041964         0.36069246         0.42639434         0.72873510         0.43270436
 factor(ag4cat)3:2  factor(ag4cat)4:1  factor(ag4cat)4:2     factor(sex)2:1     factor(sex)2:2
        1.06715164         6.22388892        10.81145980         4.02770553         1.89696556
             ald:1              ald:2              mde:1              mde:2  factor(ed4cat)2:1
        0.84892764         1.39549295         0.86957039         1.10353872         0.42868205
 factor(ed4cat)2:2  factor(ed4cat)3:1  factor(ed4cat)3:2  factor(ed4cat)4:1  factor(ed4cat)4:2
        0.52131478         0.25530364         0.39973956         0.17711483         0.29243857
factor(mar3cat)2:1 factor(mar3cat)2:2 factor(mar3cat)3:1 factor(mar3cat)3:2
        0.55438127         0.94907885         0.06175588         1.73809396

> # Note that intercepts and signs differ from other software due to differences in how reference
groups/intercepts are handled but the conclusions are similar
```

 # NOTE: MARGINAL EFFECTS EXAMPLES ARE FORTHCOMING, CHECK WEBSITE FOR UPDATES!

```
> # Alternate Bayesian approach for Section 9.2.6.1
> library(csSampling)
>
> # Part 2 of the NCS-R survey, use correct weight ncsrwtlg in filter and design set up
> ncsr$marcatc <- factor(ncsr$mar3cat, levels = 1:3, labels =c("Married", "Previously Married", "Never
Married"))
> ncsr$edcatc <-factor(ncsr$ed4cat, levels = 1:4, labels =c("0-11", "12", "13-15","16+"))
> ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
> ncsr$sexc <- relevel(ncsr$sexc, ref="Female")
> ncsr$agcatc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
> ncsr$wkstatc <- factor(ncsr$wkstat3c, levels = 1:3, labels=c("Emp", "Unemp", "NLF"))
>
> ncsr.lg <- ncsr[!is.na(ncsr$ncsrwtlg),]
>
> # Create complete data set with variables of interest
> ncsr.red <- ncsr.lg[, c("wkstatc", "marcatc", "edcatc", "sexc", "agcatc", "ald", "mde", "seclustr", "sestrat",
"ncsrwtlg")]
> ncsr.red <- ncsr.red[complete.cases(ncsr.red),]
>
> # Normalize weights to match what is done for Stan modeling
> ncsr.red$wtsc <- ncsr.red$ncsrwtlg / mean(ncsr.red$ncsrwtlg)
>
> # Survey design object
> ncsr.des <- svydesign(id = ~seclustr, strata = ~sestrat, weights = ~wtsc, nest = T, data = ncsr.red)
>
> # Bayesian approach, flat prior
> set.seed(41279)
> model_formula <- formula("wkstatc|weights(wtsc) ~ sexc + ald + mde + edcatc + agcatc + marcatc")
> mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), data = ncsr.red, family
= categorical(link = "logit"))
[1] "compiling stan model"
[1] "stan fitting"

SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
Chain 1:
Chain 1: Gradient evaluation took 0.006374 seconds
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 63.74 seconds.
Chain 1: Adjust your expectations accordingly!
Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
Chain 1:
Chain 1:  Elapsed Time: 93.945 seconds (Warm-up)
Chain 1:                84.54 seconds (Sampling)
Chain 1:                178.485 seconds (Total)
Chain 1:
```

```
[1] "gradient evaluation"
Warning message:
In mrbweights(design$cluster, design$strata, design$fpc, ...) :
  Design is sampled with replacement: only first stage used
> mod.brms$stan_fit
Inference for Stan model: anon_model.
1 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=1000.

                        mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff Rhat
b_muUnemp[1]           -1.40    0.00 0.15    -1.70    -1.50    -1.40    -1.29    -1.10  1207    1
b_muUnemp[2]           -0.22    0.01 0.40    -1.11    -0.46    -0.19     0.05     0.51  1384    1
b_muUnemp[3]           -0.15    0.00 0.19    -0.50    -0.27    -0.14    -0.02     0.23  1639    1
b_muUnemp[4]           -0.84    0.01 0.17    -1.18    -0.95    -0.84    -0.72    -0.50  1133    1
b_muUnemp[5]           -1.37    0.01 0.20    -1.75    -1.50    -1.37    -1.24    -0.97  1024    1
b_muUnemp[6]           -1.73    0.01 0.23    -2.17    -1.90    -1.74    -1.58    -1.27  1131    1
b_muUnemp[7]           -0.85    0.01 0.23    -1.33    -1.00    -0.85    -0.69    -0.40   910    1
b_muUnemp[8]           -0.84    0.01 0.25    -1.32    -1.01    -0.83    -0.68    -0.37  1072    1
b_muUnemp[9]            1.84    0.01 0.22     1.41     1.70     1.84     1.99     2.25   984    1
b_muUnemp[10]          -0.59    0.00 0.16    -0.90    -0.70    -0.59    -0.48    -0.25  1231    1
b_muUnemp[11]          -2.86    0.01 0.49    -3.96    -3.16    -2.82    -2.51    -2.01  1091    1
Intercept_muUnemp      -3.33    0.00 0.13    -3.61    -3.41    -3.32    -3.24    -3.07  1057    1
b_muNLF[1]             -0.64    0.00 0.07    -0.79    -0.69    -0.64    -0.59    -0.50  1789    1
b_muNLF[2]              0.33    0.00 0.14     0.04     0.24     0.33     0.43     0.60  1842    1
b_muNLF[3]              0.10    0.00 0.09    -0.07     0.04     0.10     0.16     0.27  1818    1
b_muNLF[4]             -0.65    0.00 0.10    -0.85    -0.72    -0.65    -0.59    -0.46   955    1
b_muNLF[5]             -0.92    0.00 0.10    -1.12    -0.99    -0.92    -0.85    -0.73  1005    1
b_muNLF[6]             -1.23    0.00 0.11    -1.43    -1.30    -1.23    -1.16    -1.01   989    1
b_muNLF[7]             -0.32    0.00 0.11    -0.53    -0.39    -0.32    -0.25    -0.13   961    1
b_muNLF[8]              0.06    0.00 0.11    -0.17    -0.01     0.06     0.14     0.27   820    1
b_muNLF[9]              2.39    0.00 0.13     2.14     2.30     2.39     2.47     2.63   851    1
b_muNLF[10]            -0.05    0.00 0.09    -0.23    -0.12    -0.05     0.01     0.14  1449    1
b_muNLF[11]             0.55    0.00 0.10     0.36     0.48     0.56     0.62     0.75  1050    1
Intercept_muNLF        -0.94    0.00 0.04    -1.01    -0.97    -0.94    -0.92    -0.87  1444    1
lprior                 -4.86    0.00 0.06    -4.98    -4.89    -4.85    -4.81    -4.74  1071    1
b_muUnemp_Intercept    -0.66    0.01 0.23    -1.08    -0.81    -0.66    -0.50    -0.21   823    1
b_muNLF_Intercept      -0.38    0.00 0.13    -0.63    -0.46    -0.38    -0.29    -0.12   929    1
lp__                -3700.72    0.16 3.50 -3708.20 -3703.03 -3700.30 -3698.18 -3695.02   502    1

Samples were drawn using NUTS(diag_e) at Sun Apr 13 13:44:47 2025.
For each parameter, n_eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor on split chains (at
convergence, Rhat=1).
```

```
> # Ordinal Regression with Russian Federation Data, Section 9.3.6
> load ("P:\\ASDA3\\Data Sets for Analysis Examples and Stata R Code\\ess6_russia.rdata")
>
> # Prepare needed outcome variable
> ess6_russia$stflife2[ess6_russia$stflife >= 0 & ess6_russia$stflife <= 1] <- 1
> ess6_russia$stflife2[ess6_russia$stflife >= 2 & ess6_russia$stflife <= 4] <- 2
> ess6_russia$stflife2[ess6_russia$stflife == 5] <- 3
> ess6_russia$stflife2[ess6_russia$stflife >= 6 & ess6_russia$stflife <= 8] <- 4
> ess6_russia$stflife2[ess6_russia$stflife >= 9 & ess6_russia$stflife <= 10] <- 5
>
> rfdata_c9 <- ess6_russia
> summary(rfdata_c9)
      idno           psu             stratify             tvtot          trstplc            vote      
 Min.   :   1   Min.   :   101.0   Length:2484        Min.   :0.000   Min.   : 0.000   Min.   :1.000  
 1st Qu.:2526   1st Qu.:   301.8   Class :character   1st Qu.:3.000   1st Qu.: 1.000   1st Qu.:1.000  
 Median :4518   Median :   412.0   Mode  :character   Median :4.000   Median : 3.000   Median :1.000  
 Mean   :4354   Mean   : 71013.6                      Mean   :4.394   Mean   : 3.486   Mean   :1.394  
 3rd Qu.:6309   3rd Qu.:   805.0                      3rd Qu.:7.000   3rd Qu.: 5.000   3rd Qu.:2.000  
 Max.   :8996   Max.   :732807.0                      Max.   :7.000   Max.   :10.000   Max.   :3.000  
                                                      NA's   :58      NA's   :94       NA's   :20     
    stflife           stfeco            gndr             agea           eisced         dweight      
 Min.   : 0.000   Min.   : 0.000   Min.   :1.000   Min.   :15.00   Min.   :1.000   Min.   :0.1561  
 1st Qu.: 4.000   1st Qu.: 2.000   1st Qu.:1.000   1st Qu.:30.00   1st Qu.:4.000   1st Qu.:0.5709  
 Median : 6.000   Median : 4.000   Median :2.000   Median :45.00   Median :5.000   Median :0.8838  
 Mean   : 5.786   Mean   : 3.806   Mean   :1.617   Mean   :45.94   Mean   :4.973   Mean   :1.0000  
 3rd Qu.: 7.000   3rd Qu.: 5.000   3rd Qu.:2.000   3rd Qu.:60.00   3rd Qu.:7.000   3rd Qu.:1.2767  
 Max.   :10.000   Max.   :10.000   Max.   :2.000   Max.   :90.00   Max.   :7.000   Max.   :4.0000  
 NA's   :26       NA's   :122                      NA's   :6                                        
    pspwght            male            agecat          marcat         stflife2    
 Min.   :0.1025   Min.   :0.0000   Min.   :1.00   Min.   :1.000   Min.   :1.000  
 1st Qu.:0.5710   1st Qu.:0.0000   1st Qu.:2.00   1st Qu.:1.000   1st Qu.:2.000  
 Median :0.8785   Median :0.0000   Median :3.00   Median :2.000   Median :4.000  
 Mean   :1.0000   Mean   :0.3829   Mean   :2.54   Mean   :1.804   Mean   :3.363  
 3rd Qu.:1.2769   3rd Qu.:1.0000   3rd Qu.:4.00   3rd Qu.:2.000   3rd Qu.:4.000  
 Max.   :4.0045   Max.   :1.0000   Max.   :4.00   Max.   :3.000   Max.   :5.000  
                                   NA's   :6      NA's   :40      NA's   :26     
>
> # Create factor variable
> rfdata_c9$marcatc <- factor(rfdata_c9$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))
> str(rfdata_c9)
'data.frame':   2484 obs. of  18 variables:
 $ idno    : int  1 2 3 4 5 6 7 8 14 17 ...
 $ psu     : int  122807 122807 122807 122807 122807 122807 122807 122807 101807 110807 ...
 $ stratify: chr  "Central FO" "Central FO" "Central FO" "Central FO" ...
 $ tvtot   : int  0 4 1 4 2 4 4 2 1 7 ...
 $ trstplc : int  1 2 2 5 3 3 2 2 NA 5 ...
 $ vote    : int  2 2 2 2 2 2 2 1 1 1 ...
 $ stflife : int  7 7 5 7 8 8 8 3 6 8 ...
 $ stfeco  : int  5 5 8 5 4 6 6 4 6 NA ...
 $ gndr    : int  1 2 2 2 2 1 2 1 2 2 ...
 $ agea    : int  22 27 23 27 23 29 30 29 36 72 ...
 $ eisced  : int  4 5 7 5 7 7 7 5 7 4 ...
 $ dweight : num  1.117 1.489 0.372 1.117 1.117 ...
 $ pspwght : num  1.177 1.381 0.401 1.035 1.204 ...
 $ male    : int  1 0 0 0 0 1 0 1 0 0 ...
 $ agecat  : int  1 1 1 1 1 1 2 1 2 4 ...
 $ marcat  : int  3 3 3 3 3 1 3 3 2 2 ...
 $ stflife2: num  4 4 3 4 4 4 4 2 4 4 ...
 $ marcatc : Factor w/ 3 levels "Married","Previous",..: 3 3 3 3 3 1 3 3 2 2 ...
```

```
 - attr(*, "written.by")= chr "Stat/Transfer v.16.4.1623.1117"
 - attr(*, "var.labels")= chr [1:16] "respondent's identification number" "primary sampling unit"
"stratification" "hours of tv avg. wkday: 0=none 1=< 0.5 hrs 2=.5-1 hrs 3=1-1.5 hrs 4=1.6-2 hrs 5=" ...
>
> rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~pspwght, data=rfdata_c9, nest=T)
>
>
> ex936 <- svymean(~factor(stflife2), design=rfsvy, na.rm=T, se=T, ci=T, keep.vars=T)
> (ex936)
                      mean     SE
factor(stflife2)1 0.045232 0.0066
factor(stflife2)2 0.200132 0.0135
factor(stflife2)3 0.209219 0.0122
factor(stflife2)4 0.424057 0.0175
factor(stflife2)5 0.121360 0.0092
>
> # Ordinal logistic using satisfaction with life, Tables 9.6 and 9.7
> summary(ex936_ordinal <- svyolr (factor(stflife2) ~ factor(agecat) + male + marcatc, design=rfsvy))
Call:
svyolr(factor(stflife2) ~ factor(agecat) + male + marcatc, design = rfsvy)

Coefficients:
                     Value Std. Error   t value
factor(agecat)2 -0.5293373 0.13614998 -3.887899
factor(agecat)3 -0.7455312 0.14332027 -5.201855
factor(agecat)4 -0.8080856 0.16561383 -4.879337
male            -0.1096234 0.09523049 -1.151138
marcatcPrevious -0.2088768 0.10542193 -1.981341
marcatcNever    -0.1371932 0.13208274 -1.038692

Intercepts:
    Value    Std. Error t value
1|2  -3.7111   0.2144   -17.3066
2|3  -1.7928   0.1668   -10.7469
3|4  -0.8348   0.1592    -5.2433
4|5   1.3841   0.1536     9.0095
(69 observations deleted due to missingness)
> #exp for odds ratio
> exp(ex936_ordinal$coef)
factor(agecat)2 factor(agecat)3 factor(agecat)4            male marcatcPrevious    marcatcNever
      0.5889952       0.4744822       0.4457105       0.8961715       0.8114952       0.8718018

> # Note test of Proptional Odds not included but there are ways to extend svyolr, see T. Lumley notes online
for details
```

```
> # Bayesian approach Example 9.3.6
> # create complete data set with variables of interest
>
> ess.red <- rfdata_c9[, c("stflife2", "agecat", "marcat", "male", "psu", "stratify", "pspwght")]
> ess.red <- ess.red[complete.cases(ess.red),]
>
> # need to normalize NCS-R weights to match what is done for Stan modeling
> ess.red$wtsc <- ess.red$pspwght / mean(ess.red$pspwght)
>
> # survey design object
> ess.des <- svydesign(id = ~psu, strata = ~stratify, weights = ~wtsc, nest = T, data = ess.red)
>
> # Bayesian approach, flat prior
> set.seed(41279)
> model_formula <- formula("stflife2|weights(wtsc) ~ factor(agecat) + factor(marcat) + male")
>
> # Run model, note the family statement needs the below code, differs from the code in the book (per Matt
Williams)
> mod.brms <- cs_sampling_brms(svydes = ess.des, brmsmod = brmsformula(model_formula),
+  data = ess.red, family = brms::cumulative(link="logit"))
[1] "compiling stan model"
[1] "stan fitting"

SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
Chain 1:
Chain 1: Gradient evaluation took 0.004907 seconds
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 49.07 seconds.
Chain 1: Adjust your expectations accordingly!
Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
Chain 1:
Chain 1:  Elapsed Time: 72.378 seconds (Warm-up)
Chain 1:                85.692 seconds (Sampling)
Chain 1:                158.07 seconds (Total)
Chain 1:
[1] "gradient evaluation"
Warning message:
In mrbweights(design$cluster, design$strata, design$fpc, ...) :
  Design is sampled with replacement: only first stage used
```

```
> mod.brms$stan_fit
Inference for Stan model: anon_model.
1 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=1000.

                   mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff Rhat
b[1]              -0.53    0.00 0.12    -0.77    -0.61    -0.54    -0.45    -0.29   682 1.00
b[2]              -0.75    0.00 0.13    -1.00    -0.84    -0.75    -0.65    -0.50   693 1.00
b[3]              -0.81    0.01 0.14    -1.09    -0.90    -0.81    -0.71    -0.55   637 1.00
b[4]              -0.21    0.00 0.10    -0.39    -0.28    -0.21    -0.14    -0.02   981 1.00
b[5]              -0.14    0.00 0.12    -0.35    -0.22    -0.13    -0.06     0.09   790 1.00
b[6]              -0.11    0.00 0.08    -0.26    -0.16    -0.11    -0.06     0.04  1273 1.00
Intercept[1]      -3.04    0.00 0.09    -3.22    -3.10    -3.04    -2.97    -2.85   680 1.00
Intercept[2]      -1.12    0.00 0.05    -1.22    -1.16    -1.12    -1.09    -1.02  1091 1.00
Intercept[3]      -0.16    0.00 0.04    -0.25    -0.19    -0.16    -0.13    -0.08  1114 1.00
Intercept[4]       2.06    0.00 0.06     1.93     2.02     2.06     2.10     2.18  1005 1.00
disc               1.00     NaN 0.00     1.00     1.00     1.00     1.00     1.00   NaN  NaN
lprior            -9.01    0.00 0.05    -9.11    -9.04    -9.01    -8.98    -8.92   882 1.00
b_Intercept[1]    -3.71    0.01 0.15    -4.01    -3.81    -3.71    -3.61    -3.43   425 1.00
b_Intercept[2]    -1.80    0.01 0.13    -2.04    -1.88    -1.79    -1.71    -1.54   487 1.00
b_Intercept[3]    -0.84    0.01 0.12    -1.07    -0.92    -0.84    -0.75    -0.60   544 1.00
b_Intercept[4]     1.38    0.01 0.13     1.12     1.30     1.39     1.47     1.61   557 1.00
lp__           -3375.75    0.10 2.09 -3380.47 -3377.07 -3375.49 -3374.16 -3372.51   454 1.01

Samples were drawn using NUTS(diag_e) at Sun Apr 13 13:57:49 2025.
For each parameter, n_eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor on split chains (at
convergence, Rhat=1).
```

```
> # Count Outcomes Section 9.4
> # HRS data for Poisson Regression using Number of Falls

> # Poisson model
> summary(hrs12)
     hhid                pn                  r11bmi          nage            nfinr           gender           secu
stratum
 Length:20554       Length:20554       Min.   : 8.9   Min.   : 20.00   Min.   :1.000   Min.   :1.000   Min.
:1.000   Min.   : 1.00
 Class :character   Class :character   1st Qu.:24.3   1st Qu.: 57.00   1st Qu.:1.000   1st Qu.:1.000   1st
Qu.:1.000   1st Qu.:17.00
 Mode  :character   Mode  :character   Median :27.5   Median : 65.00   Median :1.000   Median :2.000   Median
:1.000   Median :32.00
                                       Mean   :28.5   Mean   : 66.81   Mean   :2.237   Mean   :1.584   Mean
:1.489   Mean   :29.89
                                       3rd Qu.:31.7   3rd Qu.: 75.00   3rd Qu.:5.000   3rd Qu.:2.000   3rd
Qu.:2.000   3rd Qu.:43.00
                                       Max.   :83.0   Max.   :103.00   Max.   :5.000   Max.   :2.000   Max.
:2.000   Max.   :56.00
                                       NA's   :310    NA's   :1                        NA's   :1
     nwgthh          nwgtr          h11atota           h11itot            marcat          edcat         racecat
diabetes
 Min.   :    0   Min.   :    0   Min.   :-1510000   Min.   :      0   Min.   :1.000   Min.   :1.00   Min.
:1.000   Min.   :0.0000
 1st Qu.: 1746   1st Qu.: 1738   1st Qu.:   13962   1st Qu.:  17406   1st Qu.:1.000   1st Qu.:2.00   1st
Qu.:2.000   1st Qu.:0.0000
 Median : 3358   Median : 3479   Median :  118000   Median :  35000   Median :1.000   Median :2.00   Median
:2.000   Median :0.0000
 Mean   : 4339   Mean   : 4413   Mean   :  375979   Mean   :  61455   Mean   :1.471   Mean   :2.49   Mean
:2.119   Mean   :0.2402
 3rd Qu.: 5536   3rd Qu.: 5692   3rd Qu.:  386000   3rd Qu.:  70573   3rd Qu.:2.000   3rd Qu.:3.00   3rd
Qu.:2.000   3rd Qu.:0.0000
 Max.   :17875   Max.   :19866   Max.   :43921209   Max.   :3666240   Max.   :3.000   Max.   :4.00   Max.
:4.000   Max.   :1.0000
                                                                      NA's   :12      NA's   :100    NA's   :37
NA's   :18
   numfalls24         age65p         arthritis
 Min.   : 0.000   Min.   :0.0000   Min.   :0.0000
 1st Qu.: 0.000   1st Qu.:0.0000   1st Qu.:0.0000
 Median : 0.000   Median :1.0000   Median :1.0000
 Mean   : 1.191   Mean   :0.5223   Mean   :0.5687
 3rd Qu.: 1.000   3rd Qu.:1.0000   3rd Qu.:1.0000
 Max.   :50.000   Max.   :1.0000   Max.   :1.0000
 NA's   :9959                      NA's   :27
>
> # Create variables for example
> hrs12$nage_c=(hrs12$nage - 74.5)
> hrs12$bmi_c=(hrs12$r11bmi - 27.7)
> hrs12$male=(hrs12$gender ==1)
>
> # Create Design Object and then subset for 65 Plus
> hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
> summary(hrssvyr)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
svydesign(strata = ~stratum, id = ~secu, weights = ~nwgtr, data = hrs12,
    nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
```

```
5.034e-05 1.757e-04 2.874e-04       Inf 5.754e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23 24 25  26
27  28  29  30  31  32  33
obs        352 627 366 333 276 334 384 403 367 334 261 370 209 159 177 136 276 185 295 371 332 182 107 68 83 302
588 432 499 533 514 457 560
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2   2
2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2  2   2
2   2   2   2   2   2   2
            34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53 54  55  56
obs        483 306 542 196 370 366 669 510 489 777 691 601 687 447 297 333 685 548 153 145 56 179 152
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"     "secu"       "stratum"   
"nwgthh"     "nwgtr"     
[11] "h11atota"   "h11itot"    "marcat"     "edcat"      "racecat"    "diabetes"   "numfalls24" "age65p"    
"arthritis"  "nage_c"    
[21] "bmi_c"      "male"      
> hrssvy65 <- subset(hrssvyr, age65p==1)
> 
> # Table 9.10
> summary(ex947_poisson <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c, design=hrssvy65, 
family=quasipoisson(log)))

Call:
svyglm(formula = numfalls24 ~ male + nage_c + arthritis + diabetes + 
    bmi_c, design = hrssvy65, family = quasipoisson(log))

Survey design:
subset(hrssvyr, age65p == 1)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.635588   0.074371  -8.546 3.33e-11 ***
maleTRUE     0.257180   0.080487   3.195 0.002470 ** 
nage_c       0.014666   0.004425   3.314 0.001753 ** 
arthritis    0.736177   0.077384   9.513 1.26e-12 ***
diabetes     0.247533   0.070193   3.526 0.000938 ***
bmi_c        0.004126   0.008402   0.491 0.625615    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 8.755473)

Number of Fisher Scoring iterations: 6

Warning messages:
1: In summary.glm(g) :
  observations with zero weight not used for calculating dispersion
2: In summary.glm(glm.object) :
  observations with zero weight not used for calculating dispersion
> coefficients <- exp(coef(ex947_poisson))
> coefficients
(Intercept)    maleTRUE      nage_c   arthritis    diabetes       bmi_c 
  0.5296241   1.2932785   1.0147744   2.0879372   1.2808623   1.0041347 
> ci_exp <- exp(confint(ex947_poisson))
> ci_exp
```

```
                2.5 %    97.5 %
(Intercept) 0.4560648 0.6150479
maleTRUE    1.1000451 1.5204551
nage_c      1.0057858 1.0238433
arthritis   1.7870864 2.4394353
diabetes    1.1122679 1.4750117
bmi_c       0.9873131 1.0212430
>
> # NB model to compare to Poisson Table 9.10
> ex947_negb <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c,
+    design=hrssvy65, family=negative.binomial(theta = 1))
Warning messages:
1: In summary.glm(g) :
  observations with zero weight not used for calculating dispersion
2: In summary.glm(glm.object) :
  observations with zero weight not used for calculating dispersion
> coef <- exp(coef(ex947_negb))
> coef
(Intercept)    maleTRUE      nage_c   arthritis    diabetes       bmi_c
  0.5293307   1.2796596   1.0155015   2.0951312   1.2940333   1.0021968
> ci_exp <- exp(confint(ex947_negb))
> ci_exp
                2.5 %    97.5 %
(Intercept) 0.4525874 0.6190871
maleTRUE    1.0839844 1.5106570
nage_c      1.0065198 1.0245634
arthritis   1.7853432 2.4586728
diabetes    1.1262038 1.4868732
bmi_c       0.9860914 1.0185652
>

> # Zero Inflated Negative Binomial not easily done but with use of extensions to R tools and the csSampling
package, it can be done using a Frequentist or
> # Bayesian Method, check our website for more examples to come!
```