Homework 4

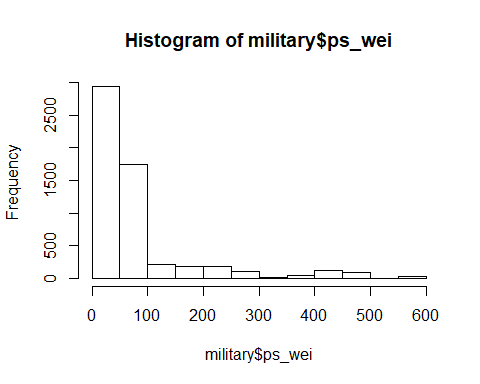
Xinyi Lin

10/5/2019

## Question a

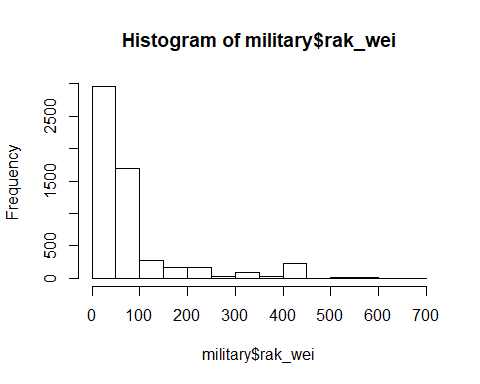
**Poststratification**

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 10.95 22.58 44.28 83.50 83.29 584.57



**Raking**

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 12.03 22.31 43.85 83.50 84.31 684.65



## Question b

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| paygrade | adj\_wei | ps\_wei | rak\_wei | pop\_wei |
| 1 | 71504.42 | 66465 | 66465 | 66465 |
| 2 | 256254.81 | 274125 | 274125 | 274125 |
| 3 | 110620.39 | 133868 | 133868 | 133868 |

According to results, we can find that even though adjusted weights of each psu given by post-stratification and raking are not the same, weights total of each payroll groups are the same and both of them are the same as population total weights in each payroll groups.

**Appendix**

knitr::opts\_chunk$set(echo = FALSE)  
library(tidyverse)  
library(survey)  
military = read.csv("./military.csv") %>%   
 mutate(XCPAY1R = ifelse(XCPAY1R==3, 2, XCPAY1R),  
 paygrade = ifelse(XCPAY1R %in% c(4,5,6,7), 3, XCPAY1R),  
 GENDER = XSEXR,  
 RACETH = XRETH4R) %>%   
 select(-XCPAY1R, -XSEXR, -XRETH4R)  
population = read.csv("./population.csv") %>%   
 mutate(GENDER = as.factor(GENDER),  
 RACETH = as.factor(RACETH),  
 paygrade = as.factor(paygrade))  
# question a  
dclus1 = svydesign(ids = ~X, strata = ~STRATUM, fpc = ~NSTRAT, weights = ~weight, data = military)  
dclus1p = postStratify(dclus1, ~ RACETH + GENDER + paygrade, population)  
military$ps\_wei = weights(dclus1p)  
summary(military$ps\_wei)  
hist(military$ps\_wei)  
pop.race = population %>%   
 group\_by(RACETH) %>%   
 summarize(Freq=sum(count\_s)) %>%   
 ungroup()   
pop.gender = population %>%   
 group\_by(GENDER) %>%   
 summarize(Freq=sum(count\_s)) %>%   
 ungroup()   
pop.pay = population %>%   
 group\_by(paygrade) %>%   
 summarize(Freq=sum(count\_s)) %>%   
 ungroup()   
rclus1 = rake(dclus1, list(~RACETH, ~GENDER, ~paygrade), list(pop.race, pop.gender, pop.pay))  
military$rak\_wei = weights(rclus1)  
summary(military$rak\_wei)  
hist(military$rak\_wei)  
# question b  
# population data  
pop\_weight = population %>%   
 group\_by(paygrade) %>%   
 summarize(pop\_wei=sum(count\_s)) %>%   
 ungroup()   
military %>%   
 group\_by(paygrade) %>%   
 summarize(adj\_wei = sum(weight),  
 ps\_wei = sum(ps\_wei),  
 rak\_wei = sum(rak\_wei)) %>%   
 ungroup() %>%   
 mutate(pop\_wei = pop\_weight$pop\_wei) %>%   
 knitr::kable()