Labour Force Participation of the Elderly in Europe - The importance of being healthy

Phi Nguyen
Julian Winkel
Claudia Guenther

Statistical Programming Languages
WS2017/2018
Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin

Outline

1. Motivation
2. Background & Methods
3. Current project status
    - Overview on quantlets
    - Highlights
4. Next steps

Labour Force Participation of Elderly in Europe

- Changing demographics in Europe: Social, health & economic challenge
  - Policy & employment implications
- SHARE panel database of micro data on health, socio-economic status and social and family networks
  - 120,000 individuals (50+), 27 participating countries
  - easySHARE as simplified data set

available on http://www.share-project.org/

Background and Methods for Quantlets

Adriaan Kalwij and Frederic Vermeulen (2005), IZA DP No. 1887

Labour Force Participation of the Elderly in Europe: The Importance of Being Healthy

- How is the labor force participation behavior of individuals aged 50-64 in 11 European countries influenced by different health indicators?
- Investigate employment potential of healthy elderly Europeans
  - Cross-sectional probit regression model (ML estimation)
  - Marginal effect at means & Wald tests
  - Counterfactual excercise
  - Lack of graphical representation

Overview on Quantlets

- Quantlet 1: Read and Clean easyshare dataset
- Quantlet 2: Summary Statistics
- Quantlet 3: Probit Regression
- Quantlet 4: Wald Test
- Quantlet 5: Counterfactual exercise
- Quantlet 6: Graphical representation

Quantlet 1: Read and Clean easyshare dataset

- Input: Data frame and wave. Output: List of data frames
- Imputation of missing values
- Conversion of country data from ISO code to a human-readable format
- Relevant error messages

```
# Split data frames into country/gender splits, then
  standardize numeric

splits     = split(df.out, f = list(df.out$country,
df.out$gender), drop = TRUE)
df.reg     = standardize.df(df.out)
df.splits = lapply(splits, standardize.df)
```

Quantlet 1: Read and Clean easyshare dataset

```
# Create necessary dummary variables for regression

 dummify = function(df) {
         df = df %>%
              dplyr::select(-country, -gender)
         model       = ~ 0 + .
         new.df      = model.matrix(model, df)
         new.df      = data.frame(new.df)
         return(new.df)
     }

df.splits = lapply(df.splits, dummify)
```

Quantlet 2: Summary Statistics

- Summary statistic tables of health characteristics of the elderly
- contains functions that calculate labor participation rates and type of labor chosen by gender and country
- Usage of ‘pmap‘ in ‘purrr‘ package to loop over groupings
- Prints tables into readable html format

Quantlet 2: Summary Statistics

	Nonparticipation	Half time	Full time
Austria	0.510	0.0513	0.439
Belgium	0.442	0.0827	0.476
Denmark	0.269	0.0782	0.653
France	0.395	0.0519	0.553
Germany	0.355	0.0534	0.591
Greece	0.299	0.1576	0.543
Italy	0.482	0.1048	0.413
Netherlands	0.337	0.0971	0.566
Spain	0.348	0.1049	0.547
Sweden	0.186	0.0898	0.725
Switzerland	0.174	0.1147	0.711
TOTAL	0.533	0.1999	0.267

Quantlet 3: Probit Regression

- Probit regression for each country and both gender groups based on ‘glm‘
- Calculation of marginal effects based on ‘probitmfx‘
- Wald test using own-built Wald test

```
1
2 allModels = lapply(df.splits, function(z){
3             z = z[-z$age50]
4         model = glm(z$labor_participationTRUE ~.,
5        family = binomial(link = "probit"), data = z)
6 return(model)
7  })
8
9 allSummaries = lapply(allModels, summary)
```


Quantlet 3: Probit Regression

```
wald.log = list()
for(i in 1:length(allSummaries)){
  SummaryElement = allSummaries[[i]]
  health = c(16:19)
  testOutput = try(joint.wald.test(allSummaries[[i]],
    health, 0.95))

  if(class(testOutput) == "try-error"){

    msg = paste0("Wald Test failed for Model Element ", i)
    warning(msg)
    wald.log[[i]] = "Error"
  } else{
    wald.log[[i]] = testOutput
  }}
```

Quantlet 4: Wald Test

- Wald test for joint significance of regression coefficients

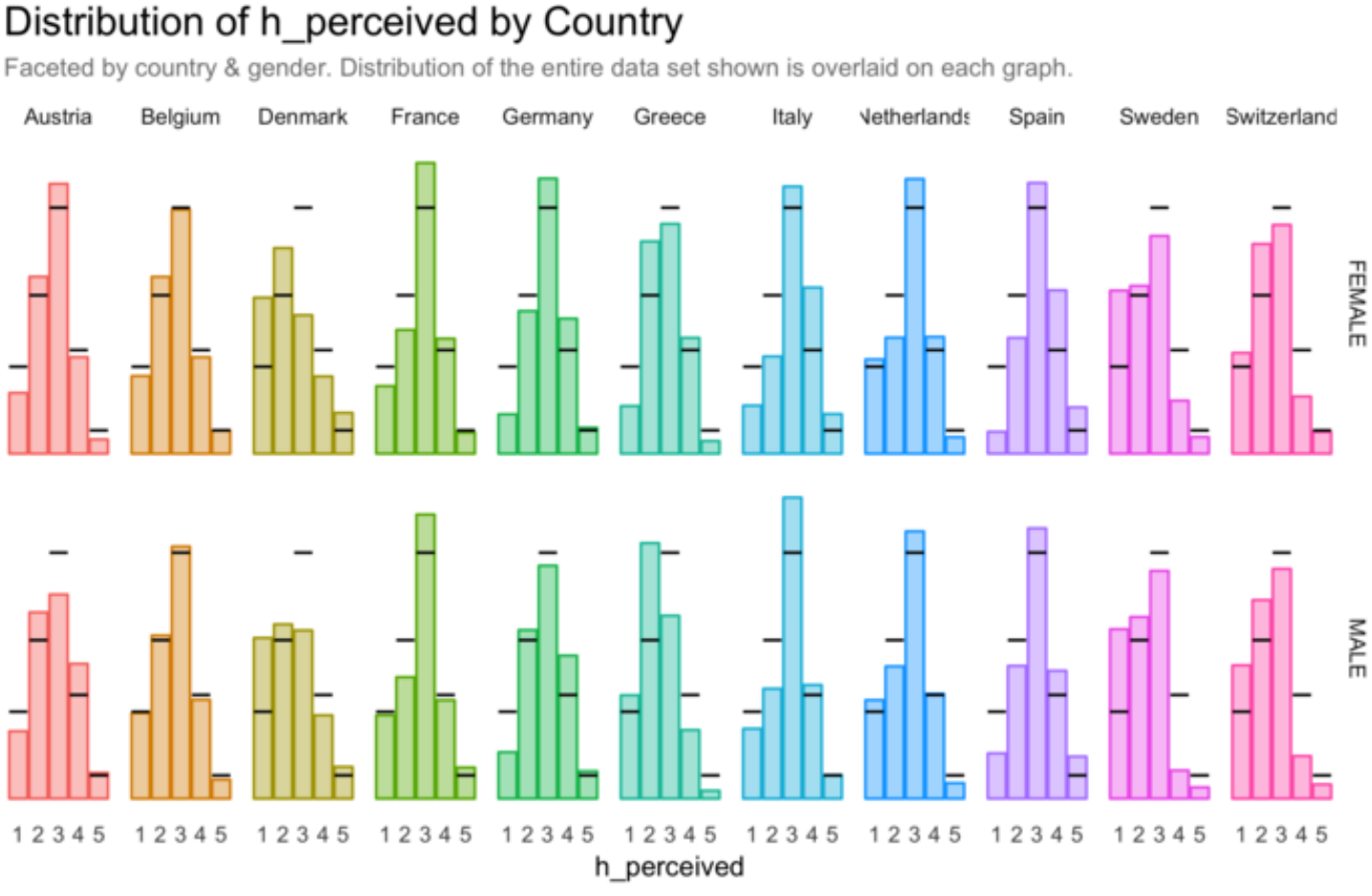
```
joint.wald.test = function(model.summary, spec, signf.l){

joint.wald.test          = numeric(6)
names(joint.wald.test)   = c("Name","W","p-value", "df")
beta                     = model.summary$coefficients[,1]
Var_beta_est             = vcov(model.summary)
W = t(beta[spec]) %*% solve(Var_beta_est[spec,spec]) %*%
  beta[spec]

   chi2                  = qchisq(signf.l, df=length(spec))
   pval                  = 1-pchisq(W,length(spec))
   joint.wald.test[1]    = "Chi2 test"
   joint.wald.test[2]    = format(   W, digits = 4)
```

Quantlet 5: Counterfactual exercise

- Estimation of current and counterfactual employment rate
- Definition of data set with perfectly healthy individuals
- Calculate decline in participation due to decline in health condition

```
X.cf    = function(model){
      X                    = model$data
      X_cf                 = X
      X_min                = data.frame(t(apply(X, 2, min)))
      names.vec            = c("h_chronic", "h_adlaTRUE", "
        h_obeseTRUE")
      X_cf[, names.vec]   = X_min[names.vec]
```

Quantlet 6: Graphical representation



Next steps for project

- Add additional graphics (choropleth maps)
- Add error messages to Wald tests
- Export all replicated results
- Review code to assure coherence with style guide
- Review code to assure coherence between quantlets