Discrete Choice, Stated Preference and Revealed Preference

Imisi Raphael Aiyetan

3/2/2021

Question 1

Excercise 16.3b and 16.3c from "A course in Evironmental Economics by Phaneuf and Requate" using R.

solution

R code:

```
# Dataframe
ic=c(5.9948,7.8305, 7.1986, 9.0011,10.483,
     5.6844,6.1743,6.8718,7.1872,7.7554,
     6.5757,8.9216,7.8807,7.9952,10.831)
oc=c(1.6558,1.378,4.3906,4.0474,1.7147,
     1.1945,1.1168,3.6592,3.9514,1.7353,
     1.7615,1.4893,4.4099,2.8629,2.0387)
#Estimate the Model
Vij=-0.6231*ic-0.4580*oc
#Find the exponential of the value function
chc=exp(Vij)
#Slice and column bind
x_3=chc[c(1:5)]
x_14=chc[c(6:10)]
x_21=chc[c(11:15)]
##question  b
prob=cbind(x_3/sum(x_3),x_14/sum(x_14),x_21/sum(x_21))
##question  c
j_5=c(10.483,7.7554, 10.831)
Redn=cbind((0.15*100*j_5)*1.36)
```

```
library(stargazer)
stargazer(t(prob), title="Predicted probabilities")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:20:26

Table 1: Predicted probabilities

0.622	0.225	0.084	0.032	0.037
0.446	0.340	0.069	0.049	0.096
0.586	0.154	0.077	0.146	0.036

```r
stargazer(Redn, title="Fifteen percent Reduction")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:20:26

Table 2: Fifteen percent Reduction

213.853
158.210
220.952

The predicted probabilities for the five choices are given in Table 1. The first, second and the third rows show probabilities for ID 3, 14 and 21 respectively. On the other hand, Table 2 reports one-time Willingnes to Pay(WTP) for a 15 percent reduction. Similarly, the first, second and the third rows show one-time WTP for ID 3, 14 and 21 respectively.

Question 2

Using the Kolkata vaccine dataset1 from "Whittington, D., Sur, D., Cook, J., Chatterjee, S., Maskery, B., Lahiri, M., … & Bhattacharya, S. K. (2009). Rethinking cholera and typhoid vaccination policies for the poor: private demand in Kolkata, India. World Development, 37(2), 399-409", estimate WTP using a parametric (probit) model for vaccine demand. Include as covariates vaccine type (chol=1 for cholera, =0 for typhoid), whether the respondent was given overnight to think about the response (ttt=1 for "time to think", =0 for no time to think), and neighborhood (til=1 for Tiljala, =0 for beliaghata). Estimate a model with only these covariates, and then a model that includes a dummy for gender (male), dummies for education levels (edu2 edu3 edu4), respondent age (age, continuous) and a dummy for whether the respondent knows someone who has 2 had the disease (KNOW).

solution

R code:

```r
library("fastDummies")
library(readr)
#load Data
Kolkata_vaccine=read_csv("Kolkata vaccine.csv")
# Create Dummmies using Fast Dummies
Kolkata= fastDummies::dummy_cols(Kolkata_vaccine)
#3a, estimate probit regression for the four model
mod=glm(yesno~priceUS+chol_cholera+ttt_TTT+til_Tiljala,
        family=binomial(link ="probit"),data=Kolkata)
mod1=glm(yesno~priceUS+chol_cholera+ttt_TTT,
        family=binomial(link ="probit"),data=Kolkata, subset = til == "Beli")
mod2=glm(yesno~priceUS+chol_cholera,
         family=binomial(link ="probit"),data=Kolkata, subset = til == "Tiljala")
mod3=glm(yesno~priceUS+chol_cholera+ttt_TTT+til_Tiljala+
           male+age+edu2+edu3+edu4+KNOW,family=binomial(link ="probit"),data=Kolkata)
```

```r
stargazer(mod, mod1, mod2,mod3, title="Probit Regressions of Vaccine demand in cholera")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:20:29

Table 3: Probit Regressions of Vaccine demand in cholera

	Dependent variable:			
	yesno			
	(1)	(2)	(3)	(4)
priceUS	-0.160^{***}	-0.156^{***}	-0.171^{***}	-0.171^{***}
	(0.012)	(0.014)	(0.023)	(0.012)
chol_cholera	-0.094	-0.125	-0.032	-0.127
	(0.096)	(0.118)	(0.166)	(0.099)
ttt_TTT	-0.294^{**}	-0.291^{**}		-0.337^{***}
	(0.119)	(0.118)		(0.122)
til_Tiljala	-0.388^{***}			-0.147
	(0.118)			(0.129)
male				-0.043
				(0.109)
age				-0.004
				(0.006)
edu2				0.557^{***}
				(0.140)
edu3				0.956^{***}
				(0.166)
edu4				0.928^{***}
				(0.198)
KNOW				0.205^{**}
				(0.100)
Constant	0.892^{***}	0.895^{***}	0.496^{***}	0.374
	(0.104)	(0.112)	(0.128)	(0.271)
Observations	826	551	275	826
Log Likelihood	-452.610	-299.755	-152.595	-430.202
Akaike Inf. Crit.	915.221	607.511	311.191	882.405
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3 shows the results of the multivariate probit regression of cholera and typhoid vaccines demand in Beliaghate and Tiljala. Column 1 and 4 report the model with and without covariates. Looking at priceUS coefficient in Column 1 and 4, It is clear that respondents are less likely to purchase if the price of a vaccine is high. In addition, the results from column 4 suggest that there is high likelihood that repondants who have primary and higher education and know someone that had disease will purchase a cholera or typhoid vaccine. Gender and age of respondents had no influence on the respondent's decision to purchase vaccines. The results from Beliaghata neighborhoods (column 2) implies that priceUS is statistically significant, which indicates that the respondents are likely to purchase the vaccines if they are given time to think about their answer. Considering column 2 and 3, it is evident that respondents from both neighborhoods have similer price sensitivity

Question 3a

Using the "vietnam choice" dataset3 from "Cook, J., Whittington, D., Canh, D. G., Johnson, F. R., & Nyamete, A. (2007). Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue, Vietnam. Economic inquiry, 45(1), 100-114.", estimate a multinomial logit model of vaccine demand ignoring (pooling together) the two TTT treatments (i.e. do not include a ttt dummy in your model – it varies at the person-level and will drop out of the choice model). The relevant variables are the following. Choice captures which of the two vaccine alternatives or the opt-out (ASC=1) was chosen. Price is continuous, three levels plus zero for the opt-out, and in USD. The vaccine's duration (dur) of exposure is effects-coded. Effects-coding centers the mean observed utility at zero, so that the observed utility for the ASC is zero. Duration dur=1 for 20 year protection, =-1 for 3 year protection and =0 for the opt-out. Vaccine type is also effects-coded, where vacc_ch=0 for the opt-out, -1 for typhoid, and =1 for cholera. Finally, the vaccine's effectiveness is captured by two effects-coded variables (eff70 and eff99). For the opt-out alternative, both eff70 and eff99=0. When effectiveness is 50%, eff70 and eff99 both =-1. When effectiveness is 70%, eff70=1 and eff99=0. When effectiveness is 99%, eff70=0 and eff99=1.

solution

R code:

```
library(mlogit)
library(readr)
#Load Data
Vietnam_choice=read_csv("Vietnam choice.csv")
```

```
## Warning: 26838 parsing failures.
##  row    col           expected    actual                 file
## 3601 time   1/0/T/F/TRUE/FALSE 10        'Vietnam choice.csv'
## 3601 w0539  1/0/T/F/TRUE/FALSE Yes       'Vietnam choice.csv'
## 3601 w0543  1/0/T/F/TRUE/FALSE Very easy 'Vietnam choice.csv'
## 3601 lntime 1/0/T/F/TRUE/FALSE 2.3025851 'Vietnam choice.csv'
## 3601 time2  1/0/T/F/TRUE/FALSE 100       'Vietnam choice.csv'
## .... ...... .................. ......... ....................
## See problems(...) for more details.
```

```
#Estimate multinomial logit regression using mlogit package for three models
Bhat=mlogit.data(Vietnam_choice, shape = "long", chid.var = "line",
                 alt.var = "alternative", choice = "choice", drop.index = TRUE)
logit=mlogit(choice~price+asc+eff70+eff99+dur+vacc_ch|-1, Bhat)
logit1=mlogit(choice~price+asc+eff70+eff99+dur+vacc_ch|-1, Bhat ,subset = nttt == "1")
logit2=mlogit(choice~price+asc+eff70+eff99+dur+vacc_ch|-1, Bhat ,subset = ttt == "1")
```

```
stargazer(logit, logit1, logit2,
          title="Conditional logit models of vaccine demand in Hue, Vietnam")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:20:34

Table 4: Conditional logit models of vaccine demand in Hue, Vietnam

	Dependent variable:		
	choice		
	(1)	(2)	(3)
price	-0.086^{***}	-0.063^{***}	-0.118^{***}
	(0.007)	(0.009)	(0.010)
asc	-0.109^{**}	-0.339^{***}	0.095
	(0.052)	(0.075)	(0.073)
eff70	0.001	-0.062	0.078
	(0.045)	(0.060)	(0.067)
eff99	0.798^{***}	0.729^{***}	0.900^{***}
	(0.045)	(0.061)	(0.068)
dur	0.044	0.036	0.047
	(0.031)	(0.042)	(0.047)
vacc_ch	0.131^{***}	0.123^{***}	0.143^{***}
	(0.029)	(0.038)	(0.044)
Observations	2,400	1,200	1,200
Log Likelihood	$-2{,}405.361$	$-1{,}225.706$	$-1{,}142.696$
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 4 reports the results of the conditional logit regression of vaccine demand in Hue, vietnam. Looking at price coefficient in Column 1 to 2, they are negative as expected and considering the absolute value, the results show that when repondents are given time to think, they are price sensitive, compared to when they have no time to think. In addition, the results show eff99 are statistically significant. This indicates that respondents are more likely to choose a 99% vaccine. However, the 70% effectiveness variable is not significant for all the three models. For 20-year duration, the insignificant coefficient of this variable indicates that repondents did not distinguish between a vaccine with 3-years duration and a vaccine with 20-years duration for TT and NT. Finally, looking at ASC coefficient, the results show that when repondents are given more time to think, they choose neither vaccines.

Question 3b

Then test whether TTT affects parameter estimates for price, asc, duration, vaccine type and effectiveness. You can do this with a set of interactions and use Wald tests to test equality of coefficients. In addition, Calculate part-worth WTP estimates for increasing duration and effectiveness among the NTT sample.

solution

R code:

```r
# Generate Willingness to pay using the estimates from logit1
coef(logit1)[- 1] / coef(logit1)[1]
```

```
##         asc        eff70       eff99         dur     vacc_ch
##   5.3633367   0.9730253 -11.5312771  -0.5733170  -1.9466461
```

```r
#Estimate multinomial logit regression question 3b
logit3=mlogit(choice~nprice+tprice+nasc+tasc+neff70+teff70+teff99+neff99
             +ndur+tdur+nvacc_ch+tvacc_ch|-1, Bhat)
```

```r
stargazer( logit3, title=" Conditional logit models of
           vaccine demand in Hue, Vietnam (Fully Interacted TTT)")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:20:35

For WTP, from Table 5, since the coefficient of 20-year duration is insignificant, then we ignore the part-worth WTP for increasing duration from 3-year to 20-year. However, the part-worth WTP for increasing effectiveness from 50% to 90% is $11.53 for NT.

Question 4a

Estimate a mixed logit model allowing the coefficient on price and the ASC to be normally distributed but other parameters as fixed coefficients. Plot the distribution of predicted coefficients (preferences) for the two random parameters (mixlbeta in Stata). You can ignore the effect of TTT for this exercise. Depending on your computer's speed, Stata may take several minutes to converge.

solution

R code:

```r
# Reshape the data for mixed logit
Bha1t=mlogit.data(Vietnam_choice, shape = "long", chid.var = "line",
                  id="respid",
                  alt.var = "alternative", choice = "choice", drop.index = TRUE)
#Estimate mixed logit
mixlogit=mlogit(choice~price+asc+eff70+eff99+dur+vacc_ch|-1,  Bha1t,
                rpar=c(price="n",asc="n"), R=100, halton=NA,print.level=0,panel = TRUE )
```

```r
stargazer(mixlogit, title="Mixed logit model of vaccine demand in Hue, Vietnam ")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 07, 2021 - 16:21:31

Table 6 reports the results of the mixed logit regression of vaccine demand in Hue, vietnam. In this model, we vary price and asc in this model and both of them are statistically significant. The absolute values of the coefficient of price, asc, eff99, dur, and vacc_ch increase and statistically significant compared to Table 3. The estimated distribution of WTP for price and asc is given as $0.80 and $0.61 respectively. The price and asc coefficents are normally distributed. Figure 3 shows the distribution of price and asc. Figure 1and 2 show the predicted individual-level coefficients for the two parameters (price and asc). See my R code for the predicted values

Question 4b

Calculate and plot individual-level expected WTP

Table 5: Conditional logit models of vaccine demand in Hue, Vietnam (Fully Interacted TTT)

	Dependent variable:
	choice
nprice	-0.063^{***}
	(0.009)
tprice	-0.118^{***}
	(0.010)
nasc	-0.339^{***}
	(0.075)
tasc	0.095
	(0.073)
neff70	-0.062
	(0.060)
teff70	0.078
	(0.067)
teff99	0.900^{***}
	(0.068)
neff99	0.729^{***}
	(0.061)
ndur	0.036
	(0.042)
tdur	0.047
	(0.047)
nvacc_ch	0.123^{***}
	(0.038)
tvacc_ch	0.143^{***}
	(0.044)
Observations	2,400
Log Likelihood	$-2,368.402$
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 6: Mixed logit model of vaccine demand in Hue, Vietnam

	Dependent variable:
	choice
price	-0.271^{***}
	(0.017)
asc	-1.133^{***}
	(0.109)
eff70	0.010
	(0.059)
eff99	1.445^{***}
	(0.071)
dur	0.166^{***}
	(0.044)
vacc_ch	0.184^{***}
	(0.039)
sd.price	0.317^{***}
	(0.022)
sd.asc	4.181^{***}
	(0.291)
Observations	2,400
Log Likelihood	$-1,694.757$
Note:	*p<0.1; **p<0.05; ***p<0.01

solution

R code:

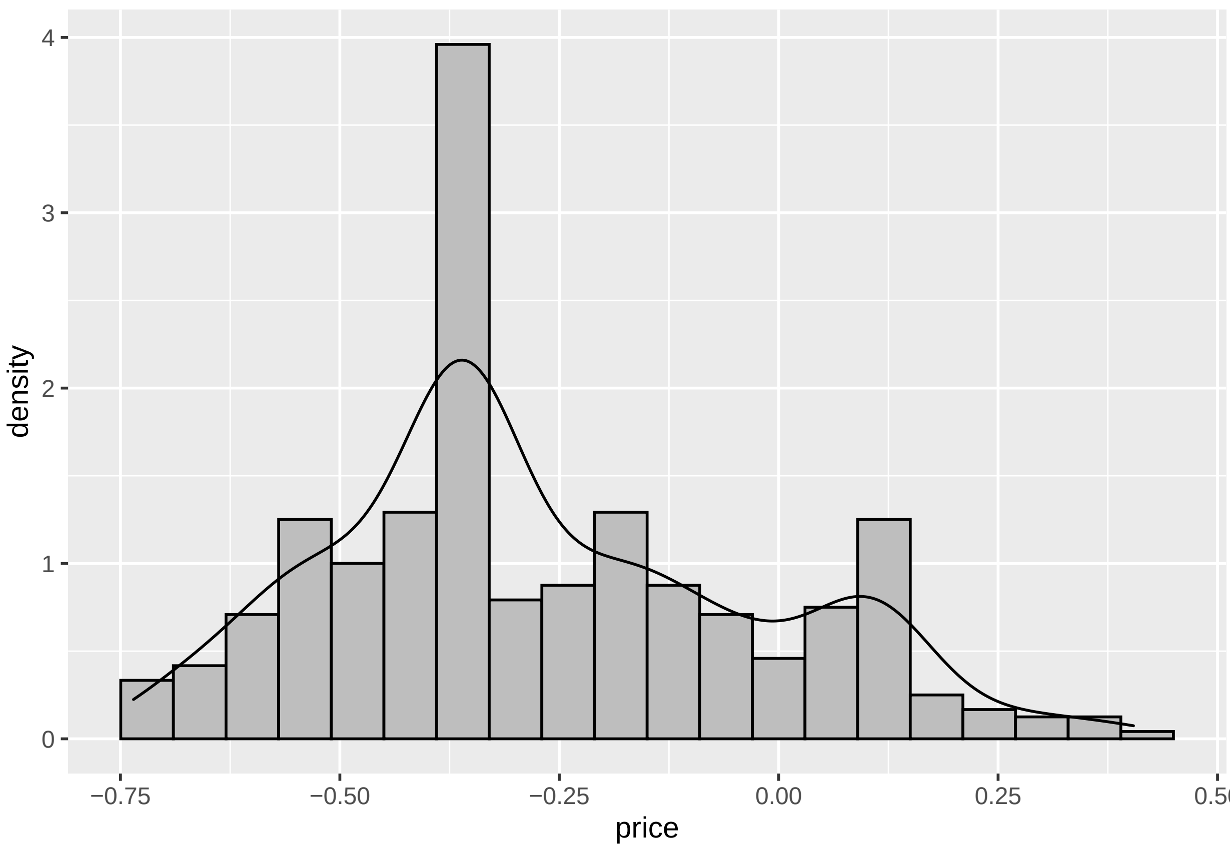
```r
# Estimate normally distribution pf fixed coefficients
pnorm(-coef(mixlogit)['price']/coef(mixlogit)['sd.price'])
```

```
##    price
## 0.804122
```

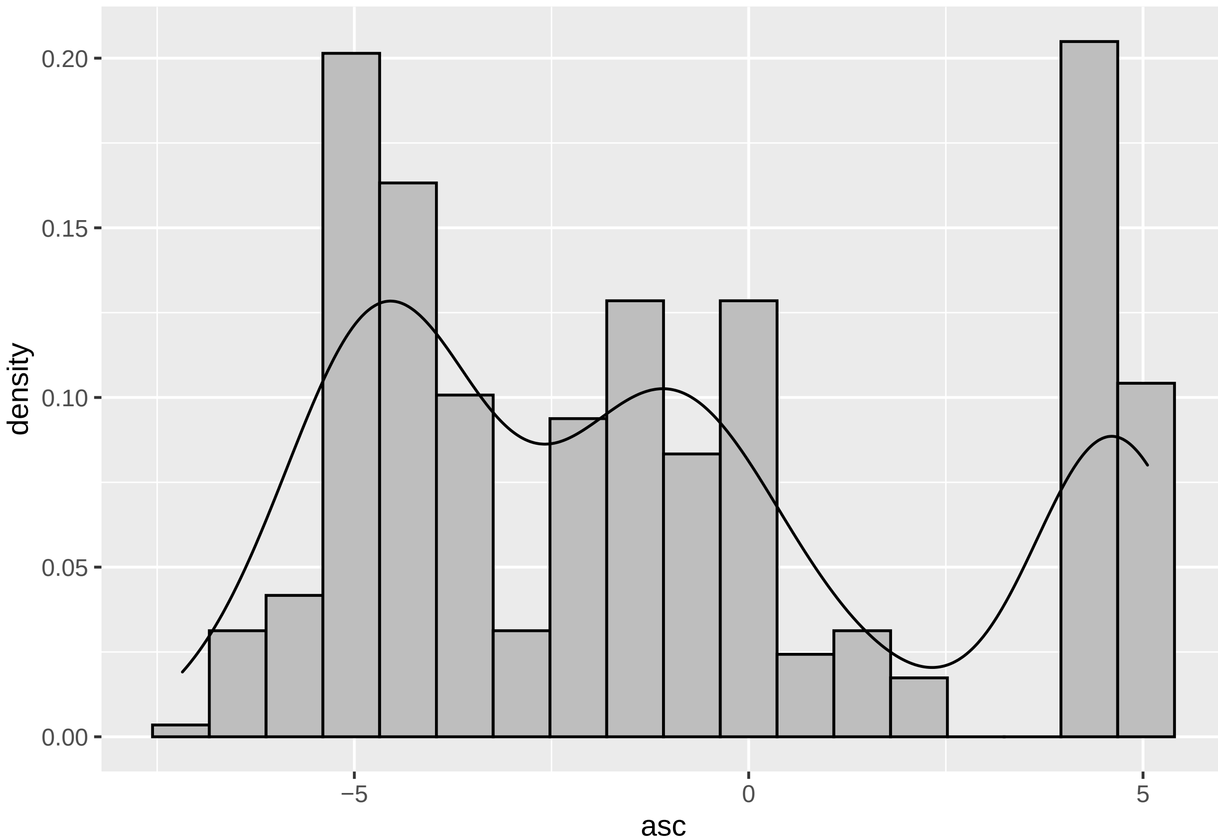
```r
pnorm(-coef(mixlogit)['asc']/coef(mixlogit)['sd.asc'])
```

```
##       asc
## 0.6068247
```

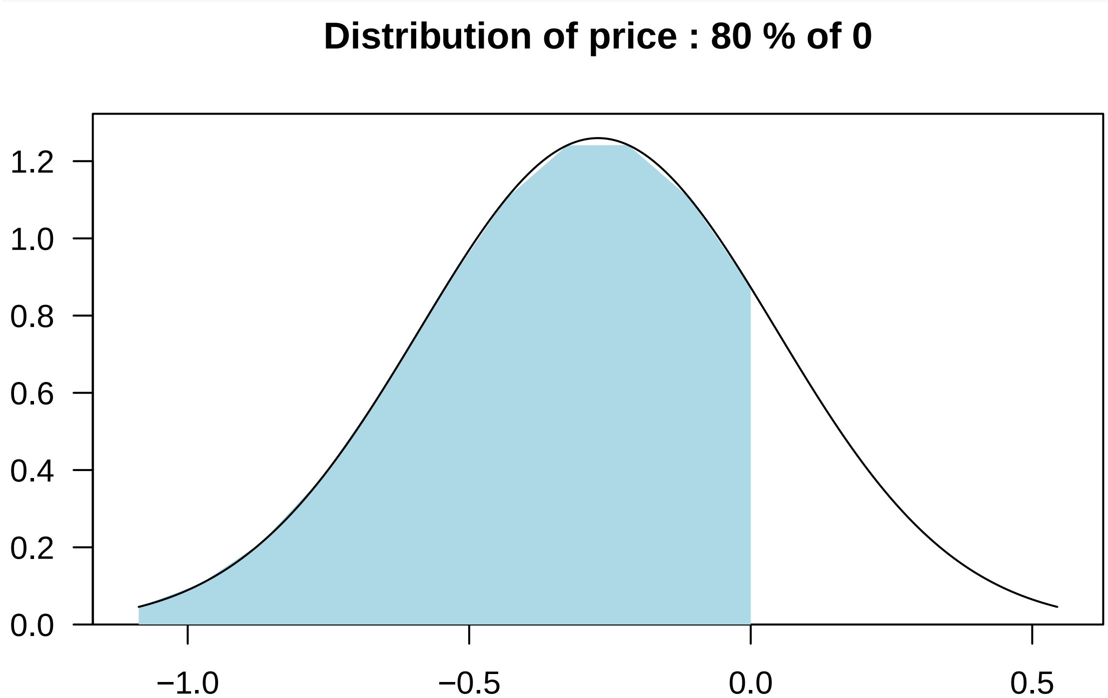
```r
#Fitted mixedlogit for individual-level expected WTP
indpar=fitted(mixlogit, type = "parameters")
#Plot of individual-level expected WTP
library("ggplot2")
ggplot(indpar) +
  geom_histogram(aes(x=price,y=..density..),
                 fill = "grey", color = "black",bins=20)+geom_density(aes(x=price,y=..density..))
```



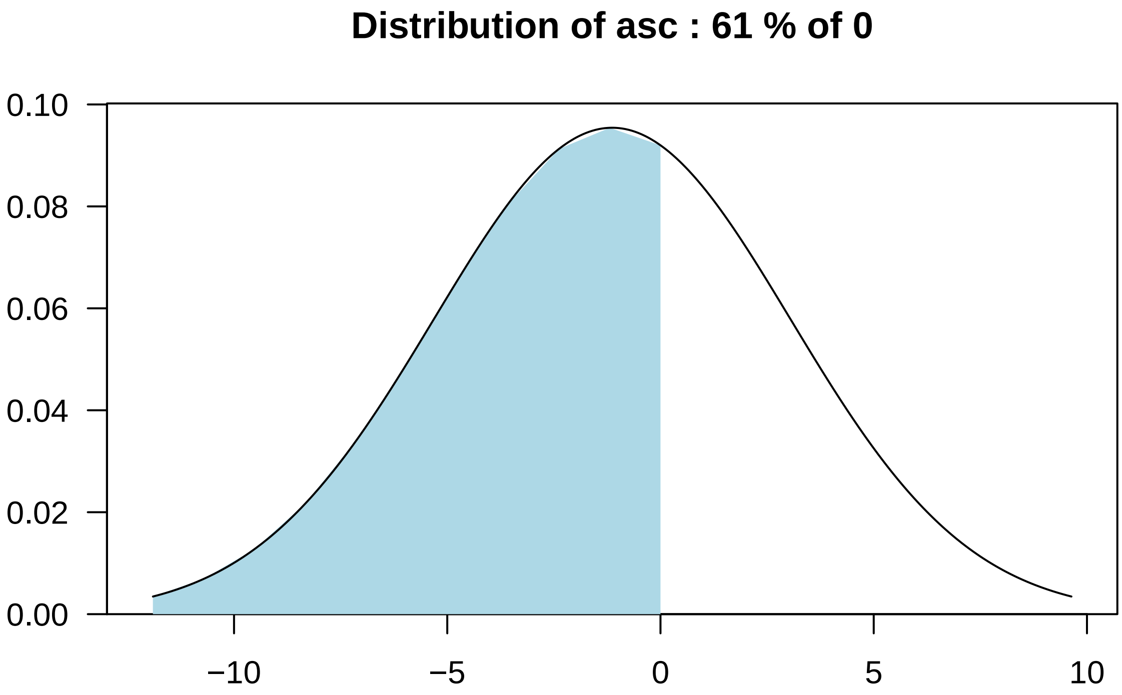

```r
ggplot(indpar) +
  geom_histogram(aes(x=asc,y=..density..),
                 fill = "grey", color = "black",bins=18)+geom_density(aes(x=asc,y=..density..))
```




```
plot(rpar(mixlogit, 'price'))
```




```
plot(rpar(mixlogit, 'asc'))
```


Distribution of asc : 61 % of 0
0.10
0.08
0.06
0.04
0.02
0.00
−10
−5
0
5
10