**SAS Homework 3**

**Group assignment**

1. Use the data in vacation.dat and run a regression with Miles (miles travelled) as the dependent variable and income, age (average age of adult members), and kids as the independent variables.
2. Run a regression model and interpret the coefficients. Comment on the model fit.

miles = -391.55+14.20\*income + 15.74\*age – 81.83 kids

All coefficients are significantly different from 0. The regression model suggests that:

When a person’s income is $1000 higher, keeping all other variables constant, that person is expected to travel 14.2 miles more per year.

When a person is 1 year older, keeping all other variables constant, that person is expected to travel 15.74 miles more per year.

When a person have 1 more kid, keeping all other variables constant, that person is expected to travel 81.83 miles less per year.

The intercept should be the miles traveled per year for observations with income=age=kids=0. However, since it’s impossible to have such observation, the intercept doesn’t mean anything on its own and is only used for prediction calculation

Model fit: The R-square is 0.3406, which suggests that the changes in 3 independent variables income, age and kids together explain 34.06% the changes in miles traveled per year.

1. Check whether there is heteroscedasticity in the model using White test.

From the Residual by Regressors graph, the variance of the residuals increases with a corresponding increase in income. We used White’s test to check for heteroscedasticity.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Heteroscedasticity Test** | | | | | |
| **Equation** | **Test** | **Statistic** | **DF** | **Pr > ChiSq** | **Variables** |
| **miles** | White's Test | 39.99 | 9 | <.0001 | Cross of all vars |

Based on the result, P-value <0.05, hence, we reject the null hypothesis and conclude that heteroscedasticity exists.

c. Run a weighted Least squares (WLS) regression. Discuss your results in a paragraph. (Comment on model fit, significance of coefficients, and the effect of doing WLS.)

After dividing all variables by income and running the WLS regression, the R-square is 0.1469, suggesting that the weighted independent variables together explain 14.69% variance in the weighted dependent variable. The coefficients are all significant and are very similar to the OLS regression coefficients we had before. After using WLS and doing the White’s test, we recognized that now heteroskedasticity doesn’t exist anymore. The P value for White’s test is 0.2672, much larger than 0.05 so we cannot reject the null hypothesis and conclude that there’s no heteroskedasticity.

2. I have provided the Sales of a durable good.

1. Using SAS and regression, estimate the Bass model. Save the regression parameters using option OUTEST. Find p, q, and M and compute peak sales and the time when that peak will occur.

p: 0.020581

q: 0.33071

time when peak sales occur (tstar): 7.90474

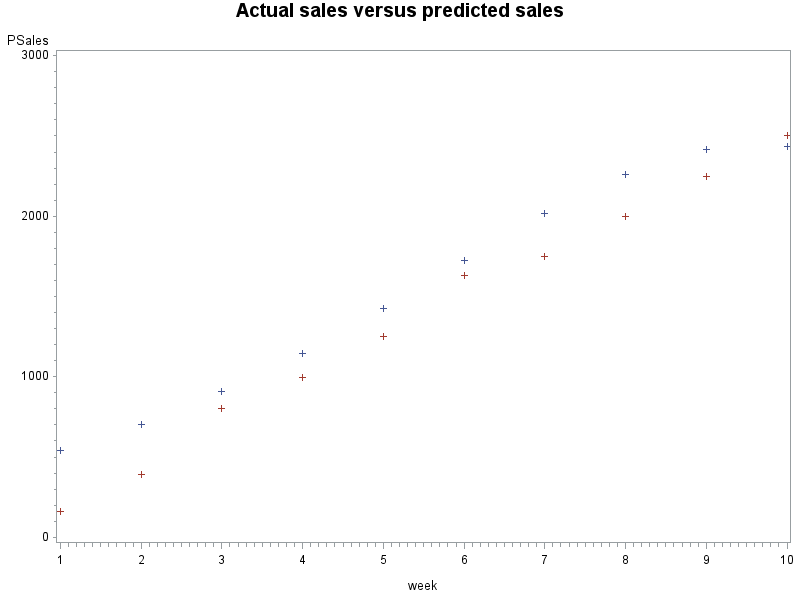
peak sales (sstar): 2446.51

| **Obs** | **\_MODEL\_** | **\_TYPE\_** | **\_DEPVAR\_** | **\_RMSE\_** | **Intercept** | **lags** | **sqrs** | **Sales** | **M** | **p** | **q** | **tstar** | **sstar** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | MODEL1 | PARMS | Sales | 125.184 | 539.731 | 0.31013 | -.000012611 | -1 | 26225.01 | 0.020581 | 0.33071 | 7.90474 | 2446.51 |

1. Predict sales in each period using only the model parameters p, q, and M and the fact that sales at time period 0=0.

|  |  |  |  |
| --- | --- | --- | --- |
| **Obs** | **week** | **Sales** | **Predicted Sales** |
| **1** | 1 | 160 | 539.74 |
| **2** | 2 | 390 | 703.45 |
| **3** | 3 | 800 | 905.80 |
| **4** | 4 | 995 | 1147.96 |
| **5** | 5 | 1250 | 1425.14 |
| **6** | 6 | 1630 | 1723.00 |
| **7** | 7 | 1750 | 2014.72 |
| **8** | 8 | 2000 | 2260.86 |
| **9** | 9 | 2250 | 2415.17 |
| **10** | 10 | 2500 | 2437.60 |

1. Plot a graph of actual versus predicted sales. (SAS code given to you in the slides)



3. A conjoint study was undertaken by a detergent manufacturer. The attributes that were considered were

Brand (Complete, Smile, Wave)

Scent (fresh, lemon, Unscented)

Whether there was a softener or not (Y, N)

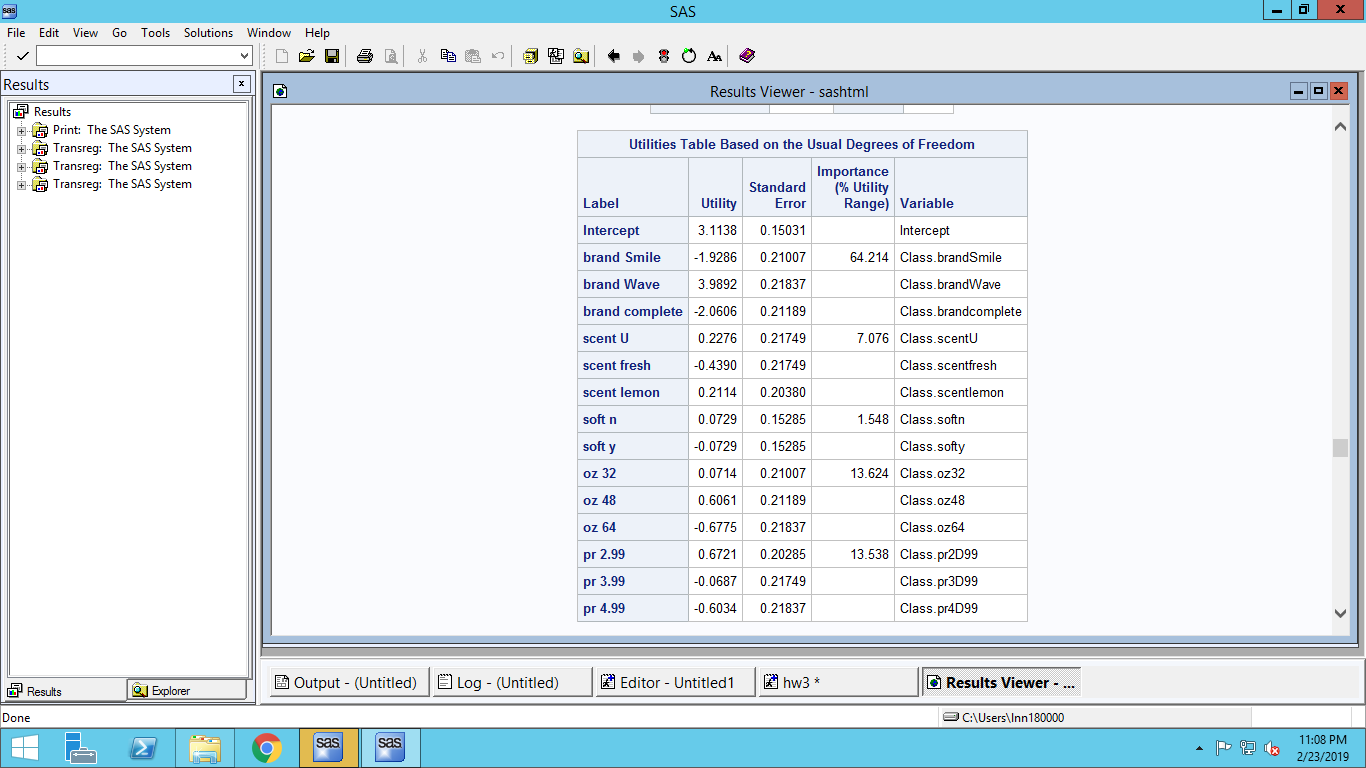
Size of packet (32, 48, 64)

Price (2.99, 3.99, 4.99)

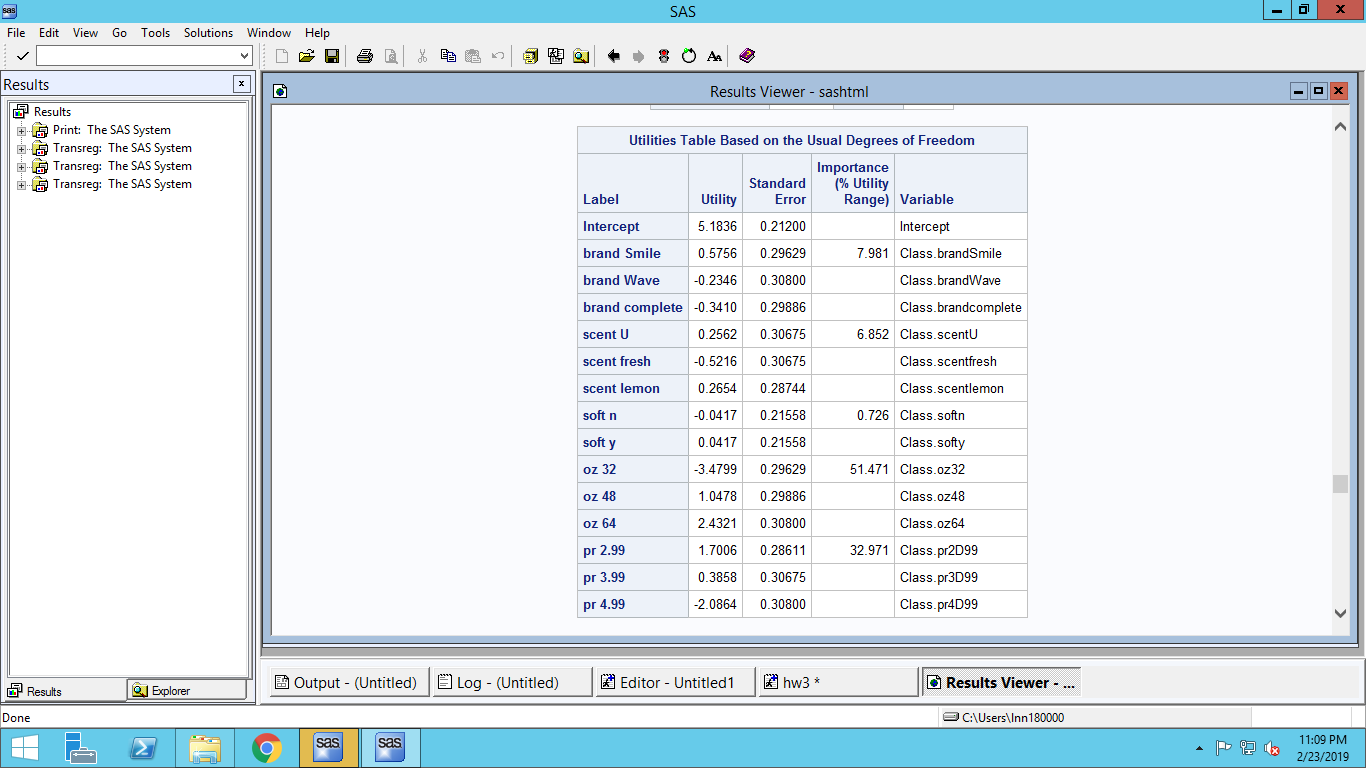
The preferences of five respondents s1, s2, s3, s4, s5 were obtained for some combination of attributes on a 1-9 point scale with 9 indicating a higher preference.

1. Find the importance weights and part-worths for each respondent using PROC TRANSREG.

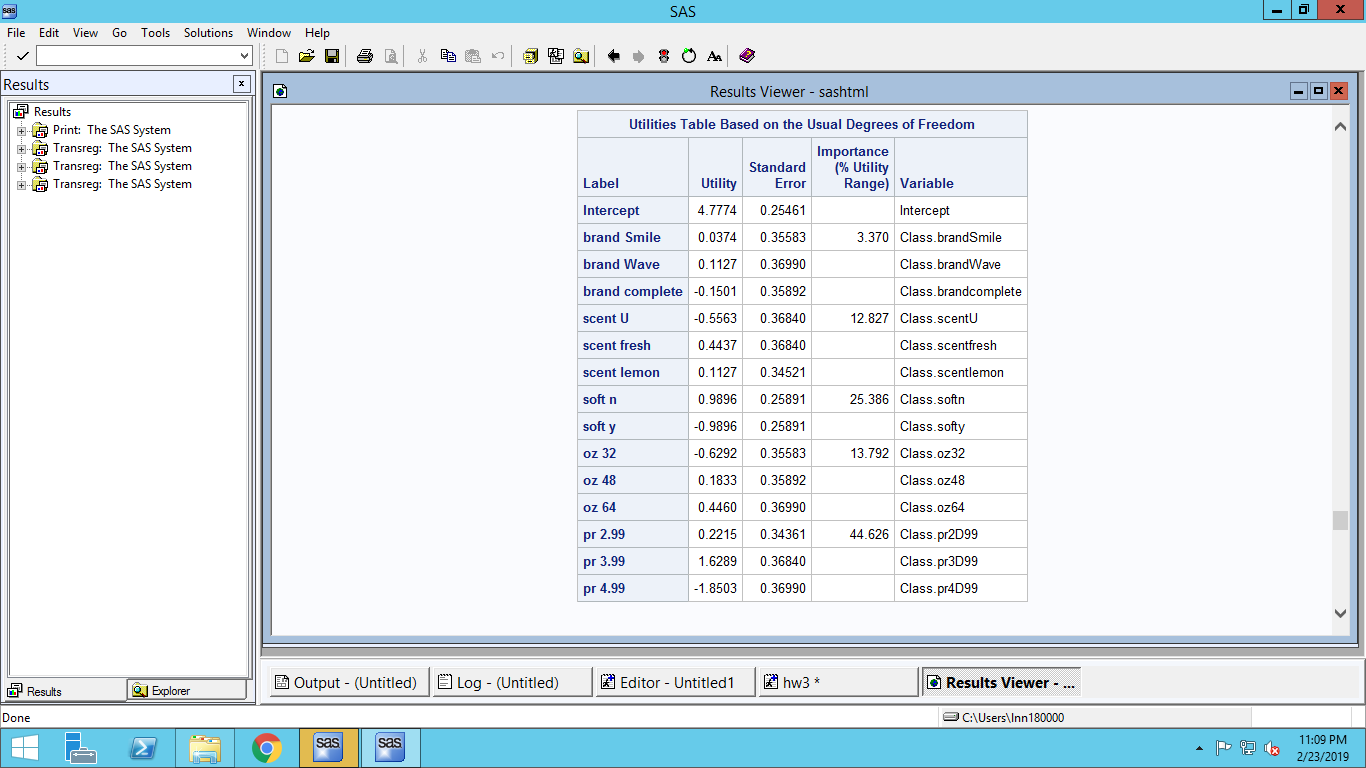
Importance weights and part-worths for person 1:



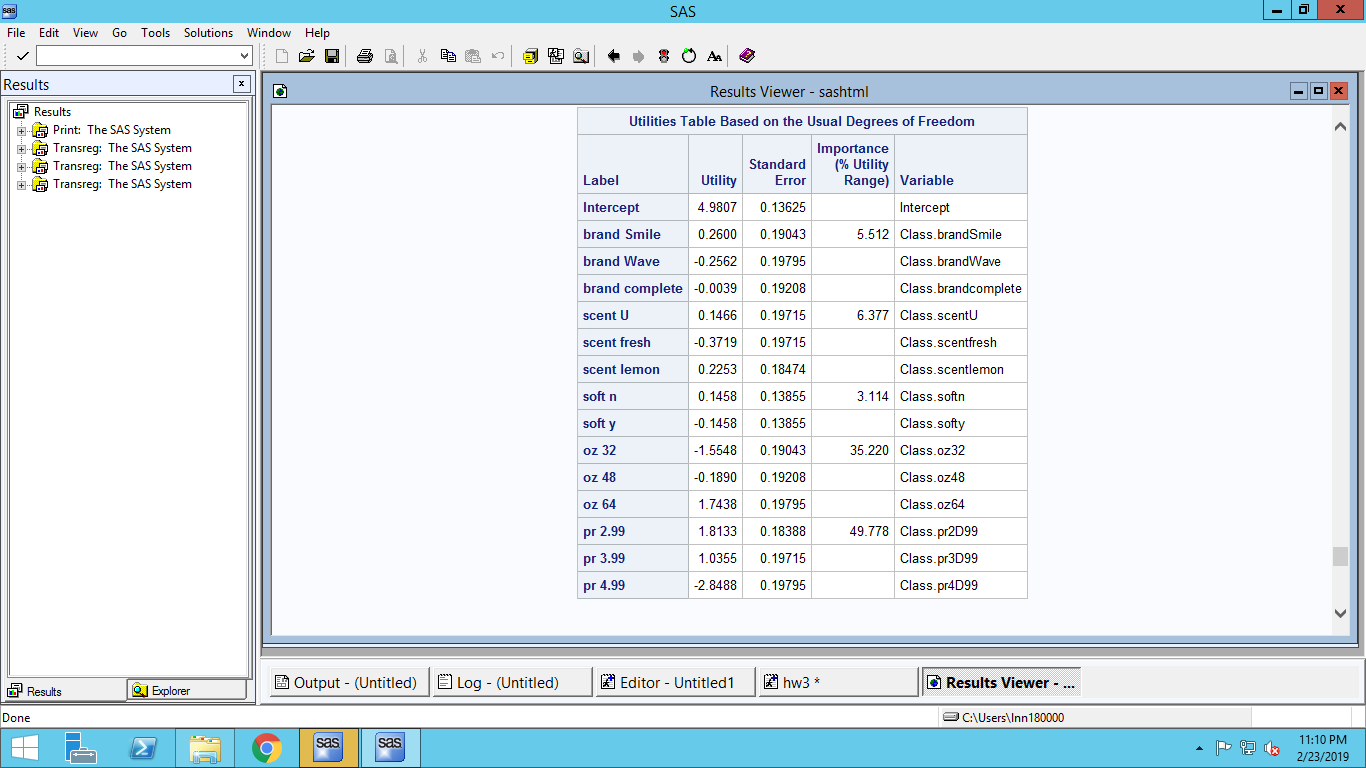
Importance weights and part-worths for person 2:



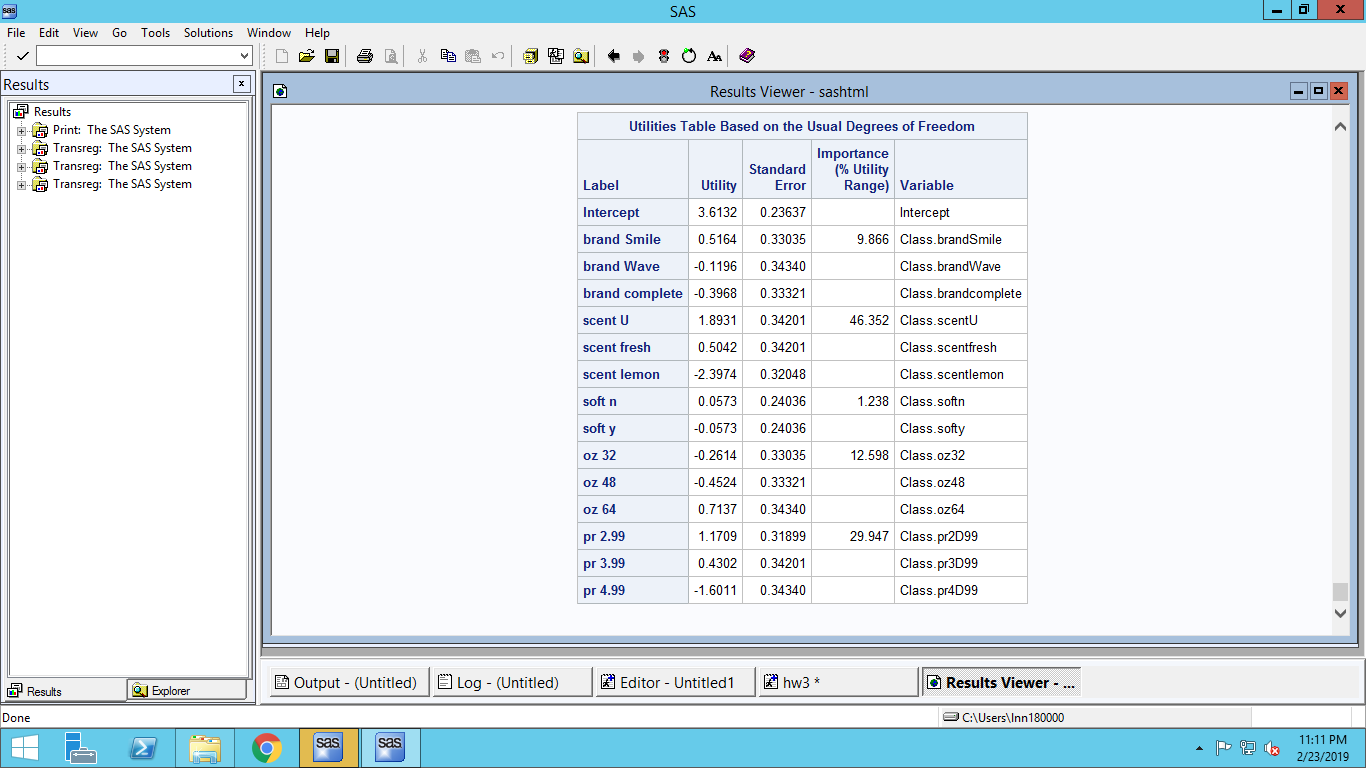
Importance weights and part-worths for person 3:



Importance weights and part-worths for person 4:



Importance weights and part-worths for person 5:



1. Predict the choice (using logit rule) for each respondent (s1-s5) for each of the following combinations using your estimates in question 1 above.

Utility of each combination for each person:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Obs** | **\_DEPVAR\_** | **U1** | **U2** | **U3** | **U4** | **U5** |
| **1** | Identity(s1) | -1.92747 | -1.16242 | -1.23650 | 5.42207 | -0.34992 |
| **2** | Identity(s2) | 4.09877 | 2.84414 | 2.30710 | 2.81173 | 3.53858 |
| **3** | Identity(s3) | -0.35957 | -0.10378 | 0.30363 | -1.02855 | 0.87539 |
| **4** | Identity(s4) | 3.63272 | 1.36651 | 1.10725 | 1.36883 | 2.17670 |
| **5** | Identity(s5) | -0.96682 | 1.68191 | 2.33005 | 2.43480 | 3.18538 |

Using logit rule, we have the following probability table:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | brand | scent | soft | oz | pr | s1 | s2 | s3 | s4 | s5 |
| 1 | complete | lemon | y | 64 | 2.99 | 0.0006 | 0.4350 | 0.1222 | 0.6575 | 0.0074 |
| 2 | Smile | fresh | y | 48 | 2.99 | 0.0014 | 0.1240 | 0.1578 | 0.0682 | 0.1041 |
| 3 | Smile | u | y | 48 | 3.99 | 0.0013 | 0.0725 | 0.2372 | 0.0526 | 0.1991 |
| 4 | Wave | u | y | 48 | 2.99 | 0.9936 | 0.1201 | 0.0625 | 0.0683 | 0.2211 |
| 5 | Smile | u | n | 48 | 2.99 | 0.0031 | 0.2484 | 0.4202 | 0.1533 | 0.4683 |

We predict that:

Person 1 is going to choose combination no. 4

Person 2 is going to choose combination no.1

Person 3 is going to choose combination no.5

Person 4 is going to choose combination no.1

Person 5 is going to choose combination no.5