# Bayesian Data Analysis - Assignment 8

November 14, 2017

## Decision analysis for the factory data

From the known information, we can get the utility function:

$$U = \begin{cases} 100 & q \geq 85 \\ -100 & q < 85 \end{cases}$$

$$E(U) = 100P(q \geq 85) - 100P(q < 85)$$

| **Machine Number** | **expected utility(euros)** |
|---|---|
| 4 | 81.75 |
| 2 | 72.40 |
| 5 | 29.15 |
| 3 | 20.90 |
| 6 | 15.60 |
| 1 | -22.40 |

Table 1: Ranking of the expected utilities

From the Table 1, we can get the following conclusion:
1. The 4th machine can make the most profit, meaning that its quality is best.
2. Besides the 4th machine, the quality of the 2nd machine is also great.
3. The 5th, 3rd, 6th machine make profits which are not high. Their qualities are acceptable.
4. The only one machine which causes loss is the 1st machine. It has the worst quality and it is not profitable. This machine should be discarded or repaired.

The expected utility of the 7th machine is 28. It is profitable, even though the expected utility is not so high. So buying a new(7th) machine is recommended.

# Appendix

## A Stan

```
data {
  int<lower=0> N; // number of data points
  int<lower=0> K; // number of groups
  int<lower=1,upper=K> x[N]; // group indicator
  vector[N] y; // measurements
}
parameters {
  real mu0;              // prior mean
  real<lower=0> sigma0; // prior std
  vector[K] mu;              // group means
  real<lower=0> sigma; // common std
}
model {
  mu~normal(mu0,sigma0);
  y ~ normal(mu[x], sigma);
}
generated quantities {
  vector[K] ypred;
  real mu7_pred;
  real y7_pred;
  mu7_pred = normal_rng(mu0,sigma0);
  for (i in 1:K)
    ypred[i]=normal_rng(mu[i],sigma);
  y7_pred = normal_rng(mu7_pred,sigma);

}
```

## B R code

```
library("rstan")
library("ggplot2")
rstan_options(auto_write = TRUE)
options(mc.cores = parallel::detectCores())

# utility function
utility<-function(sold,not_sold){
  P_sold<-sold/(sold+not_sold)
  P_not <- not_sold/(sold+not_sold)
  u <- 100 *P_sold-100*P_not
  return(u)
}

# import and organize data
raw_data<-read.table("factory.txt")
factory_data<-list(N = ncol(raw_data)*nrow(raw_data),
                   K = ncol(raw_data),
                   x = rep(1:ncol(raw_data),nrow(raw_data)),
                   y = c(t(raw_data[,1:ncol(raw_data)]))
)

# Stan
hf_fit<-stan(file="hierarchical_factory.stan",data=factory_data)
hf_result<-extract(hf_fit,permuted=TRUE)

# Calculate the utility of the six machines
U<-c()
for (i in 1:ncol(raw_data)){
  sold <-sum(hf_result$ypred[,i]>=85)
  not_sold<-sum(hf_result$ypred[,i]<85)
  u <-utility(sold,not_sold)
  U<-c(U,u)
}
print(U)

# Calculate the utility of the 7th machine
sold_7<-sum(hf_result$y7_pred>=85)
not_sold_7<-sum(hf_result$y7_pred<85)
```

```
u7<-utility(sold_7,not_sold_7)
print(u7)
```