R:

Factor

Mice library : model based imputation

#generate random variable

Rnorm, dnorm, pnorm (返回cdf的值), qnorm

#dchisq, rchisq

#rlnorm (log之后是normal distribution)

Plot(density(log(rlnorm(n=1000))))

用ANOVA => categorical variable selection

For model, need to impute the missing value first.

Lending club

Risk adjusted interest rate : risk vs interest 正相关。

不同等级的loan对应的interest rate 也是不同的。