Advanced Analytics of Finance

Problem Set 1

Elina Harutyunyan
elinah@R@mit.edu

1)

a) consistent but biased - estimator for a variance

$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$ where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

$E(\hat{\sigma}^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2$

since by LLN $\bar{X} \xrightarrow{P} E(X)$

and $\frac{1}{n} \sum_{i=1}^{n} X_i^2 \xrightarrow{P} E(X^2)$

$\bar{X}^2 \xrightarrow{P} E(X)^2$

since square is a monotonic function

by CMT & LLN

$\left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}^2\right) \xrightarrow{P} E(X^2) - E(X)^2 = \underline{Var(X)}$

→ unbiased but inconsistent

$T_n(X) = X_n$ for iid sample $(X_1 \ldots X_n)$

since from the same underlying distribution

$\Rightarrow E[T_n(X)] = E(X)$

however it does not converge to any value

b) <u>False</u> $cov(A,B)>0$
$cov(B,C)>0$

example — flip a coin 2 times

A: { First coin flip gives H }
B: { win 5 \$ }
C: { second coin flip gives H }

- if one coin is H → then win 5\$ ⇒ A, C are not correlated, hence $cov(A,C) \not> 0$ but $cov(A,B)$ & $cov(B,C)>0$.

c) $P(\text{prize in } 2 \,/\, 3 \text{ open}) = \dfrac{P(\text{prize in } 2 \text{ \& } 3 \text{ open})}{P(3 \text{ open})} =$

$$= \frac{\overbrace{P(3 \text{ open} / \text{prize in } 2)}^{=1} \, \overbrace{P(\text{prize in } 2)}^{\frac{1}{3}}}{P(3 \text{ open})}$$

$P(3 \text{ open}) = P(3 \text{ open} \mid \text{prize in } 1) \times P(\text{prize in } 1) + P(3 \text{ open} / \text{pr in } 2) \times P(\text{pr in } 2) + P(3 \text{ open} \mid \text{pr in } 3) \times P(\text{pr in } 3) =$

$$= \frac{1}{2} \times \frac{1}{3} + 1 \times \frac{1}{3} + 0 \times \frac{1}{3} = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$$

$\Rightarrow P(\text{prize in } 2 / 3 \text{ open}) = \dfrac{1/3}{1/2} = \boxed{\dfrac{2}{3}}$

hence switching gives more chance of winning!

2]

a) $X = 1, 2, 3$

$P_R = \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}$

$$P(X=3 / X>1) = \frac{P(X=3 \,\&\, X>1)}{P(X>1)} = \frac{1/3}{2/3} = \frac{1}{2}$$

b) $Y = X + \varepsilon \sim N(0,\sigma^2)$ $\quad E[\varepsilon X] = E[\varepsilon]E(X) = 0$

$X \sim N(0,1)$ $\quad \Leftrightarrow \quad Y \sim N(0, 1+\sigma^2)$

$f(x|y) = ?$ $\quad Y|X \sim N(x, \sigma^2)$

$$f(x/y) = \frac{f(y|x) f(x)}{f(y)} = \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(y-x)^2}{\sigma^2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\frac{1}{\sqrt{2\pi}\sqrt{(1+\sigma^2)}} e^{-\frac{1}{2}\frac{y^2}{(1+\sigma^2)}}} =$$

$$= \frac{\sqrt{1+\sigma^2}}{\sqrt{2\pi}\,\sigma} \underbrace{e^{-\frac{1}{2}\left(\frac{y^2}{\sigma^2} - \frac{2xy}{\sigma^2} + \frac{x^2}{\sigma^2} + x^2 - \frac{y^2}{(1+\sigma^2)}\right)}}$$

$$e^{-\frac{1}{2}\left(\frac{\overbrace{y^2(1+\sigma^2)}^{= y^2 + \cancel{y^2\sigma^2}} - 2xy(1+\sigma^2) + \underbrace{x^2(1+\sigma^2) + x^2\sigma^2(1+\sigma^2)}_{x^2(1+\sigma^2)(1+\sigma^2)} - \cancel{y^2\sigma^2}}{\sigma^2(1+\sigma^2)}\right)} =$$

$$= e^{-\frac{1}{2}\left(\frac{x^2(1+\sigma^2)^2}{\sigma^2(1+\sigma^2)} - \frac{2xy(1+\sigma^2)}{\sigma^2(1+\sigma^2)} + \frac{y^2}{\sigma^2(1+\sigma^2)}\right)}$$

$$= \frac{\sqrt{1+\sigma^2}}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{\left(x - \frac{y}{1+\sigma^2}\right)^2}{\left(\frac{\sigma}{\sqrt{1+\sigma^2}}\right)^2}\right)} \Rightarrow f(x/y) \sim N\left(\frac{y}{1+\sigma^2}, \left(\frac{\sigma}{\sqrt{1+\sigma^2}}\right)^2\right)$$

c) $X_t \sim N(0,1)$ iid over time

observe if $X_t > 0$.

density of X, conditional on $X_t > 0$

if $\frac{f(x)}{P(X \geq 0)}$ → since st. Normal $= \frac{1}{2}$

$\Rightarrow$ conditional $f(X / X > 0) = 2f(x)$

hence $E(X_{t+1} \mid X > 0) = \int_0^\infty 2x\, f(x)\,dx =$

$= 2 \int_0^\infty x \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^\infty x e^{-\frac{x^2}{2}} dx =$

$= \frac{2}{\sqrt{2\pi}} \left. -e^{\frac{-x^2}{2}} \right|_0^\infty = \frac{2}{\sqrt{2\pi}}$

d) $X_{t+1} = m_t + \varepsilon_{t+1}$ $\quad m_t = E_t(X_{t+1}) = E(X_{t+1} / t)$

$E_t(\varepsilon_{t+1}) = 0$

$Var(X_{t+1}) \overset{?}{=} Var(m_t) + E(Var_t(\varepsilon_{t+1}))$

law of total variance: $Var(Y) = E[Var(Y/X)] + Var(E[Y/X])$

hence we have! * conditioning on time t

$Var(X_{t+1}) = E(Var_t(X_{t+1})) + Var(E_t(X_{t+1}))$

$Var_t(X_{t+1}) = Var_t(\varepsilon_{t+1})$ since $m_t = E_t(X_{t+1}) =$ constant

$E_t(X_{t+1}) = m_t$

hence $Var(X_{t+1}) = Var(m_t) + E[Var_t(\varepsilon_{t+1})]$

3) $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ $\quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ $\quad E(x_i^2) = \sigma^2+\mu^2$

$E(s^2) \overset{?}{=} \sigma^2 \qquad E(\bar{x}) = \mu$

$E(s^2) = \frac{1}{n-1} E\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2\right) =$

$= \frac{1}{n-1}\left(n \underbrace{E(x_i^2)}_{\sigma^2+\mu^2} - 2E(\bar{x})E\left(\sum_{i=1}^{n} x_i\right) + nE(\bar{x}^2)\right) =$

$=$ since $\bar{x} = \frac{\sum x_i}{n} \Rightarrow E(\bar{x}^2) = \frac{E(\sum x_i)^2}{n^2}$

$2E(\bar{x})E\left(\sum_{i=1}^{n} x_i\right) = \frac{2}{n} E\left(\left(\sum_{i=1}^{n} x_i\right)^2\right) \Rightarrow \frac{2}{\cancel{n}} n^2 E(\bar{x}^2)$

$= \frac{1}{n-1}\left(n\sigma^2 + n\mu^2 \underbrace{- 2nE(\bar{X}^2) + nE(\bar{X}^2)}_{-nE(\bar{x}^2)}\right)$

$\boxed{nE(\bar{x}^2)} = nE\left(\frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2\right) = \frac{1}{n} E\left[\left(\sum_{i=1}^{n} x_i\right)^2\right] =$

$= \frac{1}{n} E\left[\sum_{i=1}^{n} x_i^2 + \sum_{j\neq i} x_i x_j\right] =$

iid

$= \frac{1}{n} \underbrace{\sum_{i=1}^{n} E(x_i^2)}_{(\sigma^2+\mu^2)n} + \frac{n^2-n}{n} \underbrace{E(x_i)E(x_j)}_{\mu^2} = \sigma^2+\mu^2+n\mu^2-\mu^2 = \boxed{\sigma^2+n\mu^2}$

so $E(s^2) = \frac{1}{n-1}\left(n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2}\right) = \frac{\sigma^2\cancel{(n-1)}}{\cancel{(n-1)}} = \sigma^2$

posterior of R^e_{t+1}

$$p(R^e_{t+1}/x) = \int_{\mu} p(R^e_{t+1}/\mu)\, p(\mu/x)\, d\mu =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(R^e_{t+1}-\mu)^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi}\, V_T} e^{-\frac{(\mu-m_T)^2}{2V_T^2}}\, d\mu$$

$$= \frac{1}{2\pi\sigma V_T} \int_{-\infty}^{\infty} \underbrace{e^{-\frac{(R^e_{t+1}-\mu)^2}{2\sigma^2}-\frac{(\mu-m_T)^2}{2V_T^2}}}\, d\mu$$

$$e^{\left(-\frac{V_T^2(R^e_{t+1}-\mu)^2+\sigma^2(\mu-m_T)^2}{2\sigma^2 V_T^2}\right)} =$$

$$= e^{-\left(\frac{V_T^2(R^{e\,2}_{t+1} - 2\mu R^e_{t+1} + \mu^2) + \sigma^2(\mu^2 - 2\mu m_T + m_T^2)}{2\sigma^2 V_T^2}\right)} =$$

$$= e^{-\left(\frac{(V_T^2+\sigma^2)\mu^2 - 2\mu(V_T^2 R^e_{t+1} + \sigma^2 m_T) + (V_T^2 R^{e\,2}_{t+1} + \sigma^2 m_T^2)}{2\sigma^2 V_T^2}\right)} =$$

$$= e^{-\left(\frac{\left(\mu - \frac{R^e_{t+1} V_T^2 + m_T\sigma^2}{V_T^2+\sigma^2}\right)^2 + \frac{V_T^2 R^{e\,2}_{t+1} + \sigma^2 m_T^2}{V_T^2+\sigma^2} - \left(\frac{R^e_{t+1} V_T^2 + m_T\sigma^2}{V_T^2+\sigma^2}\right)^2}{2\sigma^2 V_T^2 / (V_T^2+\sigma^2)}\right)}$$

we separate out the terms not dependent on μ

$$\Rightarrow p(R_{t+1}^e \mid X) = \frac{1}{2\pi \sigma V_T} e^{\frac{(-R_{t+1}^e - m_T)^2}{2(V_T^2+\sigma^2)}} \int_{-\infty} e^{-\frac{\left(\mu - \frac{R_{t+1}^e V_T^2 + m_T \sigma^2}{V_T+\sigma^2}\right)^2}{2\sigma^2 V_T^2/(V_T^2+\sigma^2)}} d\mu$$

From here we eventually obtain that

$$p(R_{t+1}^e \mid X) = \frac{1}{\sqrt{2\pi(V_T^2+\sigma^2)}} e^{\left(-\frac{(R_{t+1}^e - m_T)^2}{2(V_T^2+\sigma^2)}\right)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} B} e^{\left(\frac{-(\mu-A)^2}{2B^2}\right)} d\mu$$

where

$$A = \frac{R_{t+1}^e V_T^2 + m_T \sigma^2}{V_T^2+\sigma^2} \quad B = \frac{\sigma^2 V_T^2}{V_T^2+\sigma^2}$$

hence integrates to 1

$$\Rightarrow p(R_{t+1}^e \mid X) = \frac{1}{\sqrt{2\pi(R_T^2+\sigma^2)}} e^{\left(-\frac{(R_{t+1}^e - m_T)^2}{2(V_T^2+\sigma^2)}\right)}$$

hence $R_{t+1}^e \sim N(\bar{\mu}_{R,t}, \bar{\sigma}_{Rt}^2)$

where

$\bar{\mu}_{R,t} = m_T$

& $\bar{\sigma}_{Rt}^2 = V_T^2 + \sigma^2$.

Question 4

g) $E_t[w(1+R_{t+1}) + (1-w)(1+R_{f,t})] - \frac{\alpha}{2} Var_t[w(1+R_{t+1}) + (1-w)(1+R_{f,t})] = E_t[1+R_{f,t} + wR^e_{t+1}] - \frac{\alpha}{2} Var_t[1+R_{f,t} + wR^e_{t+1}]$

$= 1 + E_t(R_{f,t}) + wE_t(R^e_{t+1}) - \frac{\alpha}{2}[Var_t(R_{f,t}) + \underbrace{Var(wR^e_{t+1})}_{w^2 var(R^e_{t+1})} + 2 cov_t(R_{f,t}, wR^e_{t+1})]$

~~$E_t[R_{f,t} \times wR^e_{t+1}]$~~ here we assume that market excess returns are indep. of the risk free rate & hence $cov_t(R_{f,t}, wR^e_{t+1}) = 0$

a month before → so interest rates are iid over time

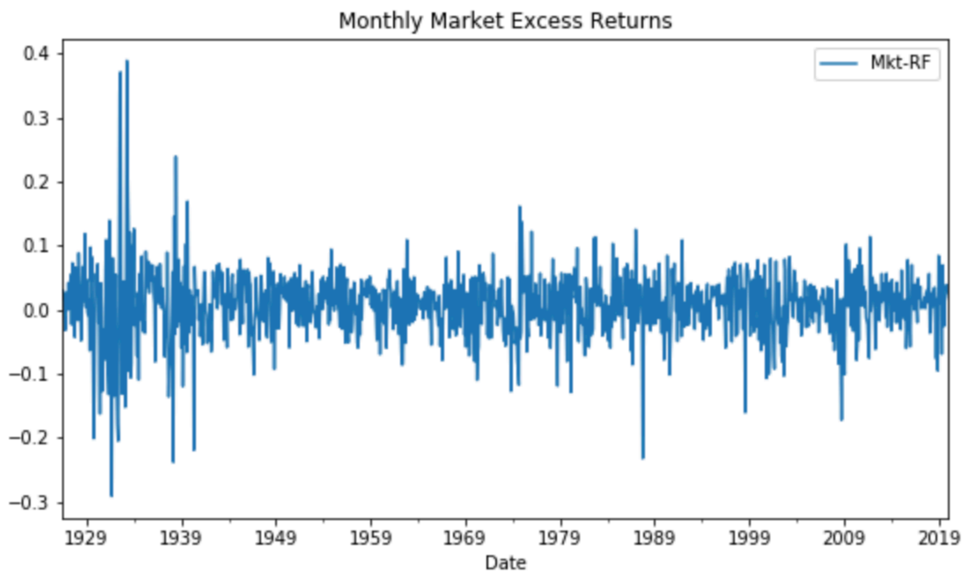
hence $\frac{\partial}{\partial w} = E_t(R^e_{t+1}) - \alpha w\, var(R^e_{t+1}) = 0$

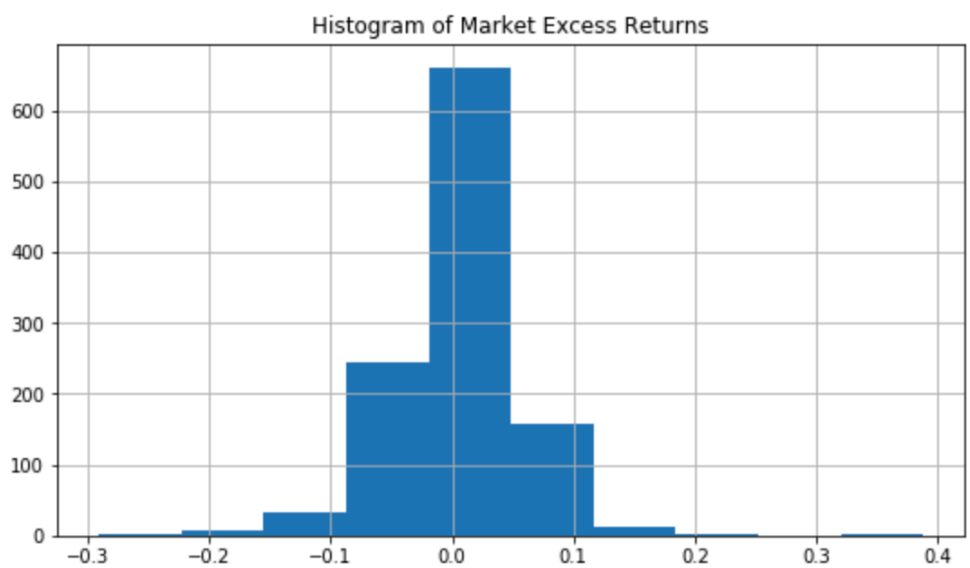
$$w = \frac{E_t(R^e_{t+1})}{\alpha\, var(R^e_{t+1})} = \frac{\mu_{R,t}}{\alpha \sigma^2_{R,t}}$$

Advanced Analytics of Finance
Problem Set 1
Elina Harutyunyan
elinahar@mit.edu

Problem 4: Developing a market timing strategy

a)





b) The market excess returns are distributed normally with the following parameters:

$$r_t^e \sim \mathcal{N}(\mu, \sigma^2)$$

The sigma is known and set to be:

$$\sigma = 5.4\%$$

The prior distribution of the mean is normal with the following parameters:

$$\mu \sim N(m_0, v_0^2)$$

I take the hyperparameters to be:

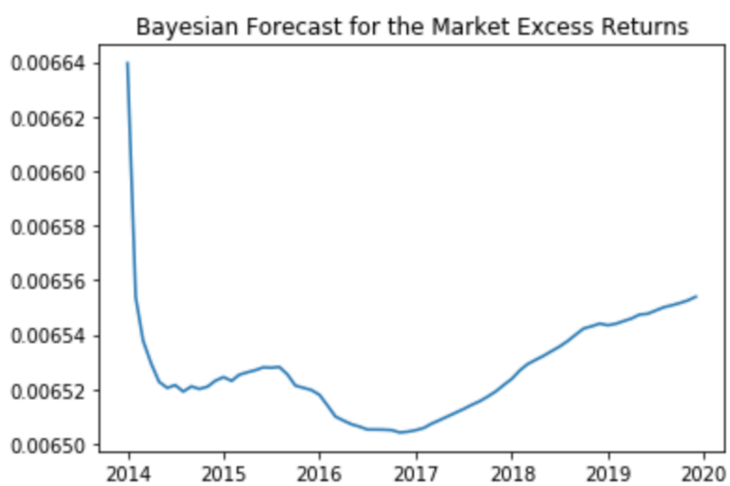
$$m_0 = 5\%, v_0^2 = 0.09\%$$

The parameters of the posterior distribution of the mean equal:

$$m_T = \frac{m_0\sigma^2 + \bar{R}Tv_0^2}{\sigma^2 + Tv_0^2}, \; v_T^2 = \frac{v_0^2\sigma^2}{\sigma^2 + Tv_0^2}$$

From here we obtain the one month ahead posterior for mean to equal 0.66395%.

d) For MSE we obtain very low value of 0.00116.



e) Assuming that market excess returns evolve over time according to this model:

$$r_{t+1}^e = a_0 + a_1 r_t^e + \varepsilon_{t+1}$$

Where

$$\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_e^2)$$

We perform OLS regressions on the lagged market excess returns and get coefficients:

$$a_0 = 0.0057, a_1 = 0.1136$$

Refer to the table below for other relevant statistics.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Mkt-RF   R-squared:                       0.013
Model:                            OLS   Adj. R-squared:                  0.012
Method:                 Least Squares   F-statistic:                     13.68
Date:                Thu, 20 Feb 2020   Prob (F-statistic):           0.000228
Time:                        00:44:13   Log-Likelihood:                 1574.6
No. Observations:                1049   AIC:                            -3145.
Df Residuals:                    1047   BIC:                            -3135.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0057      0.002      3.424      0.001       0.002       0.009
lag_mktrf      0.1136      0.031      3.699      0.000       0.053       0.174
==============================================================================
Omnibus:                      164.024   Durbin-Watson:                   1.994
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2150.233
Skew:                           0.209   Prob(JB):                         0.00
Kurtosis:                      10.001   Cond. No.                         18.4
==============================================================================
```

f) Since the error term is normally distributed with mean 0 and variance σ_e^2 it follows that the excess market return is also normally distributed with following parameters:

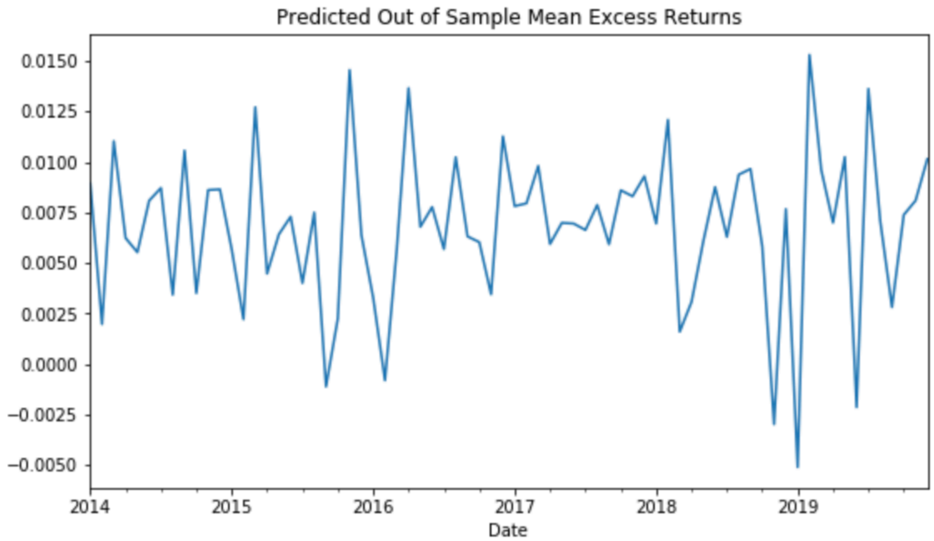
$$E_t(r_{t+1}^e) = a_0 + a_1 E_t(r_t^e)$$

Hence at time t, the excess market return r_t^e is realized, it is a constant, hence the expectation of one-month ahead market excess return is

$$\hat{\mu}_{r,t} = a_0 + a_1 r_t^e$$

The variance of one-month ahead market excess returns are therefore

$$\hat{\sigma}^2{}_{r,t} = Var_t(r_{t+1}^e) = Var_t(\varepsilon_{t+1}) = \sigma_e^2$$



The graph above shows out of sample mean excess returns predicted by the OLS model described before. Out of sample MSE of the forecasts equals 0.00124, which is very close to the MSE produced by the model before, however slightly higher. Therefore, the OLS performs slightly better at predicting excess returns of the market.

h) Below you can see the plots of the monthly NAV of the portfolio with weights calculated according to the formula:

$$\omega_t = \frac{\mu_{r,t}}{\alpha\sigma^2_{r,t}},$$

Where

$$\mu_{r,t} = \mathrm{E}_t[r^e_{t+1}] \text{ and } \sigma^2_{r,t} = \mathrm{Var}_t[r^e_{t+1}]$$

The graphs of both strategy returns show that updating our posterior distribution of the one-month ahead market excess returns every month gives us higher returns than just using coefficients obtained from the OLS regression in the training set. This holds when the risk aversion parameter equals 1 or 10.