FINM 36700 **Mark Hendricks**
Autumn 2022

Homework #3B

1 Modeling Volatility and VaR

Use the returns on the S&P 500 (r^M) and 1-month T-bills, (r^f) provided in
"barnstable_analysis_data.xlsx".

For the full sample of SPY returns, 1926-2022, calculate the LEVEL, (not log,) excess market returns (against the treasury returns.) We use this level excess return data throughout this section.

1. Historic VaR.

   Starting at $t = 61$, calculate the historic-based VaR, based on the expanding sample from period 1 (Jan 1926) to $t-1$. By historic VaR, we mean simply taking the 5th quantile for the historic sample up to time $t-1$. Of course, a sample size that is not a multiple of 100 will require some interpolation to get a 5th quantile. Your statistical package should handle this fine.

   Denote this as $\tilde{r}^{\text{VaR},.05}_{t,\text{historic}}$, which is the best estimate of the time-t VaR based on data through $t-1$.

   (a) Plot $\tilde{r}^{\text{VaR},.05}_{t,\text{historic}}$ over time.

   (b) Calculate the frequency of periods in which $\tilde{r}_t < \tilde{r}^{\text{VaR},.05}_{t,\text{historic}}$.

   (c) What drawbacks do you see in this historic VaR?

2. Volatility

   We will calculate a time-series of volatility estimates using a few different methods. For each, we use σ_t to denote our estimate of the time-t return volatility, as based on data over periods 1 (Jan 1926) through $t-1$, but not including t itself.

   - Expanding Series[1]

   $$\sigma^2_{t,\text{expanding}} = \frac{1}{t-1}\sum_{\tau=1}^{t-1}\tilde{r}^2_\tau$$

     Begin the calculation at $t = 61$, so that the first estimate is based on 60 data points.

   - Rolling Window

   $$\sigma^2_{t,\text{rolling}} = \frac{1}{m}\sum_{l=1}^{m}\tilde{r}^2_{t-l}$$

     Use $m = 60$, and begin the calculation at the $t = 61$, (so that the calculation has a full 60 data points.)

[1] Note that here we use a slightly biased version of the usual variance estimator by ignoring μ and dividing by the number of data points.

(a) For each of these three methods, plot σ_t. (Plot the vol, not the variance.)

(b) For each method, calculate the 5th percentile, 1-month-ahead VaR. We use a slight simplification of the normal VaR formula, by dropping μ from that formula, and rounding the normal distribution z-score to -1.65.

$$\tilde{r}^{\text{VaR},0.05}_{t,\text{vol}} = -1.65\sigma_t$$

(c) For each of these three vol-based VaR estimates, calculate the frequency of periods in which $\tilde{r}_t < \tilde{r}^{\text{VaR},.05}_{t,\text{vol}}$.

(d) Compare and contrast your results among each other and relative to the historic method in the previous problem.

3. CVaR

Re-do the previous two problems, but this time calculating CVaR instead of VaR, (still for $q = .05$.) That is, calculate CVaR for

- the empirical cdf
- the normal model, using expanding volatility estimates
- the normal model, using rolling volatility estimates

4. **Extra**

We may discuss this after Midterm 1, but it is not part of the Midterm 1 material.

For the VaR calculations of Problem 2.2, and 2.3 try using the following models to estimate volatility:

- Exponentially Weighted Moving Average (EWMA)

$$\sigma^2_{t,\text{EWMA}} = \theta\sigma^2_{t-1,\text{EWMA}} + (1-\theta)\tilde{r}^2_{t-1}$$

  Rather than estimating θ, simply use $\theta = 0.97$, and initialize with $\sigma_1 = 0.15$.

- GARCH(1,1) model.

  To estimate GARCH(1,1), try using the ARCH package in Python. The default estimation implementation is fine.

  You should be familiar with EWMA and GARCH from the August Review.

2 Barnstable's Analysis

The Risk of Stocks in the Long-Run: The Barnstable College Endowment [HBS 9-296073].

This section is not graded, and you do not need to submit your answers.

- But you are expected to consider these issues and be ready to discuss them.
- This section requires no empirical analysis; answer solely based on the material given in the case.

1. Barnstable's Philosophy
    (a) What has Barnstable's investment strategy been in the past?
    (b) Explain the logic behind their view that stocks are safer in the long run.
    (c) What assumptions underly Barnstable's belief in the long-run safety of stocks?

2. Two Proposals
    (a) Describe the two proposals Barnstable is considering for how to take advantage of their view regarding the long-run safety of stocks.
    (b) How is the trust different from simply shorting the risk-free rate to take a levered position in stocks?
    (c) Do these proposals take the same bet on long-run stock performance? In what outcomes will they have different returns?
    (d) Do the two proposals differ in their risk?

3. Do you recommend a direct investment in the S&P, the trust or the puts?

3 Estimating Underperformance

Use the returns on the S&P 500 (r^M) and 1-month T-bills, (r^f) provided in **"barnstable_analysis_data.xlsx"**.

Barnstable's estimates of mean and volatility are based on the subsample of 1965 to 1999. We consider this subsample, as well as 2000-2022, as well as the full sample of 1926-2022. We only have data through August of 2022, but no adjustment is needed for the fact that you have only the partial year–just use what you have.

1. Summary Statistics

   (a) Report the following (annualized) statistics.

		1965-1999		2000-2022		1926-2022	
		mean	vol	mean	vol	mean	vol
levels	r^M						
	$\tilde{r}^M$						
	r^f						
logs	r^M						
	$\tilde{\mathrm{r}}^M$						
	r^f						

   (b) Comment on how the full-sample return stats compare to the sub-sample stats.
   Comment on how the level stats compare to the log stats.

2. Recall the following...

   - If $x \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$, then

$$\Pr\left[x < \ell\right] = \Phi_{\mathcal{N}}\left(L\right)$$
$$L = \frac{c - \mu_x}{\sigma_x}$$

     where $\Phi_{\mathcal{N}}$ denotes the standard normal cdf.

   - Remember that cumulative log returns are simply the sum of the single-period log returns,

$$\mathrm{r}^M_{t,t+h} \equiv \sum_{i=1}^{h} \mathrm{r}^M_{t+i}$$

   - It will be convenient to use and denote sample averages. We use the following notation for an h-period average ending at time $t+h$,

$$\bar{\mathrm{r}}^M_{t,t+h} = \frac{1}{h}\sum_{i=1}^{h} \mathrm{r}^M_{t+i}$$

Calculate the probability that the cumulative market return will fall short of the cumulative risk-free return.[2]

$$\Pr\left[R^M_{t,t+h} < R^f_{t,t+h}\right] \tag{1}$$

To analyze this analytically, convert the probability statement above to a probability statement about mean log returns.

(a) Calculate (1) using the subsample 1965-1999.

(b) Report the precise probability for $h = 15$ and $h = 30$ years.

(c) Plot the probability as a function of the investment horizon, h, for $0 < h \leq 30$ years.

3. Use the sample 1965-2022 to reconsider the 30-year probability. As of the end of 2022, calculate the probability of the stock return underperforming the risk-free rate over the next 30 years. That is, $R^M_{t,t+h}$ underperforming $R^f_{t,t+h}$ for $0 < h \leq 30$.

4. Let's consider how things turned out relative to Barnstable's 1999 expectations.

   (a) What was the probability (based on the 1999 estimate of μ,) that the 23-year market return, $R^M_{t,t+23}$, would be smaller than that realized in 2000-2022? Note that we are asking about the market return, not the excess market return. Continue using the 1965-1999 sample standard deviation for σ.

   (b) Suppose Barnstable had implemented the put-option strategy with the growing strike of 6%? Based on the 2000-2022 performance, what is the probability that the 2000-2029 cumulative market return will be low enough to make the puts in the money? For the calculation, update your estimates of μ, σ to be the estimates based on 1965-2022.

[2] Note that this is essentially what Barnstable is estimating, but with the 1965-1999 sample estimate that the risk-free rate was 6% during that period.

4 Extensions

This section is not graded, and you do not need to submit your answers. We may discuss some of these extensions.

1. Let's consider how Barnstable would have done if they had positioned the problem in excess returns rather than beating a fixed benchmark based on historic risk-free rates.

   (a) Based on the end-of-1999 parameters, $\tilde{\mu}$ and $\tilde{\sigma}$, what was the probability that the log market excess return, $\tilde{\mathrm{r}}^M$ would have a smaller 22-year average than what happened in 2000-2022, based on $\tilde{\mu}$ estimated from the 1965-1999 average??

   $$p(h) = \Pr\left[\bar{\tilde{\mathrm{r}}}^M_{t,t+21} < \bar{\tilde{\mathrm{r}}}^M_{2000,2022}\right]$$

   (b) Suppose they struck the puts to the realized risk-free rate instead of a fixed 6%. Calculate the probability that the puts will be in the money, given what has already occurred in 2000-2022. As above, calculate this using full-sample parameter estimates of 1965-2022.

   (c) Explain the risk Barnstable took on the future movements in the risk-free rate.

2. The probability of shortfall calculations for Barnstable presume that we can simply use sample estimates for the probability calculations that involve population parameters for the means and variances of log returns. Let's consider how this estimation uncertainty leads to uncertainty in the probability of shortfall calculation.

   (a) Assuming log returns are normally distributed, (or just using the CLT,) calculate the 5th and 95th percentiles for our confidence interval of the mean of excess log returns: $\tilde{\mu} = \mu - \mu_f$.

   (b) Denote these $\tilde{\mu}_{.05}, \tilde{\mu}_{.95}$, and recalculate the probability of shortfall from (1). Plot the confidence bounds as a function of the investment horizon, $0 < h \leq 30$.

   (c) Has $\tilde{\mu}$ from 2000-2022 been outside the 5th and 95th percentile estimates derived from the data in 1965-1999?

   (d) Does uncertainty around $\tilde{\mu}$ materially impact the estimates? Does it change your recommendation to Barnstable regarding the trust versus the puts?

3. Does the analysis change much if we use the pre-1965 data in our estimates of the market return and risk-free rate?

4. The case assumes stock returns are iid. and lognormally distributed.

   $$\mathrm{r}^M_t \sim \mathcal{N}\left(\mu, \sigma^2\right), \forall t$$
   $$r^M_t = \exp\left(\mathrm{r}^M_t\right) - 1$$

   The following table lists return stats given by the case,

   Notice that Barnstable's mean estimate is for the level-return, r^M_t, while Barnstable's volatility estimate corresponds to the log-return, r^M_t. Here and elsewhere it is at times useful to translate mean and volatility from level to logs—at least for the lognormal case.

Barnstable case estimates

	mean	**variance**
level return of stocks	$\mathbb{E}\left[r_t^M\right] = 0.13$	$\text{var}\left[r_t^M\right] =?$
log return of stocks	$\mathbb{E}\left[\mathrm{r}_t^M\right] =?$	$\text{var}\left[\mathrm{r}_t^M\right] = (0.16)^2$
log riskless benchmark	$\mathbb{E}\left[\mathrm{r}^f\right] = 0.06$	

Use the following formulas to translate log-return means and variances to level-return means and varianaces.

$$\mathbb{E}\left[r_t^M\right] = e^{\mu+\frac{1}{2}\sigma^2} - 1$$
$$\text{var}\left[r_t^M\right] = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

(a) Use Barnstable's estimates in the table above, along with these equations to calculate the mean log return, μ.

Do you get the same answer as calculated by the case, 0.117? (In footnote 3.)

(b) Also calculate $\text{var}\left[r_t^M\right]$.

(c) Are the mean and volatility bigger for level or log returns?