Solutions Final Exam Advanced Risk Management, 26 March, 2019

1. (a) Advantage: choosing m large means that the sample percentile is always based on many observations, and hence is more precise. Disadvantage: the HS VaR assumes that the past m days are representative for the next day; when m is very large, this is unrealistic: we are averaging over other regimes than the current regime.

   (b) A very large negative return $R_t = -4\sigma$ means that the $VaR_{t+1}^{0.01}$ goes up, from the old maximum of $\{-R_{t-j}\}_{j=1}^{100}$ to the new value $VaR_{t+1}^{0.01} = -R_t = 4\sigma$. A very large positive return $R_t = 4\sigma$, on the other hand, will have no effect on $VaR_{t+1}^{0.01}$, because the first percentile (in the left tail) is not affected.

   (c) A positive and significant t-ratio implies that the exceedances are significantly positively correlated. This means that if $I_t = 1$ today, then the probability of $I_{t+1} = 1$ will be larger than on other days, when $I_t = 0$. The exceedances will be clustered, so the probability of being hit by two large negative returns on consecutive days becomes larger than $p^2 = 10^{-4}$, which would be the case if the exceedances are independent. This creates a problem, because the bank's economic capital may quickly become insufficient to absorb the shocks, when it is hit by such a cluster of negative shocks.

   (d) The problem is that HS is not flexible enough to adapt smoothly to changes in volatility and hence risk. A simple extension that may solve this is to use Weighted Historical Simulation, where percentiles are computed based on weights $\lambda^j/(1-\lambda)$ for some $\lambda \in (0,1)$ given to R_{t-j} (so recent returns are given a higher weight than returns a longer time ago). A similar result may be obtained by using the RiskMetrics approach (EWMA conditional variance with normal percentiles) or Filtered Historical Simulation, where HS is applied to the standardised shocks from a GARCH or EWMA model.

   (e) The main drawback of VaR is that it does not offer information on how bad the losses can be on days when the loss exceeds the VaR. ES is the mean loss conditional on it being larger than the VaR, and therefore does provide this information. A theoretical advantage of ES is that, in contrast to VaR, it is guaranteed to be sub-additive (the risk of a sum is less than or equal to the sum of the risks).

2. (a) The clearest indication of volatility clustering is the coefficient α, which indicates that the conditional variance tomorrow will go up when today $(R_t - \mu)^2$ is very large, i.e., the return is far from the mean. Because the t-ratio of $\hat{\alpha}$ is approximately 7.5, much larger than the critical value of 1.96, we find that this effect is significant. The significance of $\hat{\beta}$ may also be used as argument for volatility clustering, but only in connection with $\hat{\alpha} > 0$; if $\beta > 0$ but $\alpha = 0$, then there is mean-reversion in the variance, but not volatility clustering.

   (b) In the GARCH(1,1) model, the unconditional variance is given by $\mathrm{Var}(R_t) = \omega/(1-\alpha-\beta)$, provided that $\alpha + \beta < 1$. In this case, we find $\hat{\alpha} + \hat{\beta} = 0.0299 + 0.9671 = 0.9970 < 1$ and $\hat{\omega} = 1.1635 \times 10^{-3}$, so the estimate of the unconditional variance is $1.1635 \times 10^{-3}/(1 - 0.9970) = 0.3878$.

   (c) We clearly see that the histogram has a higher peak than the normal density, and that it is a bit lower than the normal density for z around ± 1. Although we do not see very extreme outcomes, the fact that the horizontal axis runs from -6 to 4 shows that there were realisations in this interval. All this indicates a larger kurtosis than the normal distribution, which would mean in particular that the VaR with a probability of 1% or less would be underestimated by a method based on $z_t \sim N(0,1)$; so the results do not support such a method.

   (d) First, we note that, because $x_t = -z_t = (-R_{t+1} + \mu)/\sigma_{t+1}$, we have

$$0.01 = \Pr(R_{t+1} < -VaR_{t+1}^{0.01}) = \Pr\left(x_{t+1} > \frac{VaR_{t+1}^{0.01} + \mu}{\sigma_{t+1}}\right).$$

In combination with the Pareto distribution given in the question, this means that

$$0.01 = 0.05 \left(\frac{VaR_{t+1}^{0.01} + \mu}{u\sigma_{t+1}} \right)^{-1/\xi},$$

which may be solved for $VaR_{t+1}^{0.01}$:

$$VaR_{t+1}^{0.01} = -\mu + \sigma_{t+1} u \left(\frac{0.01}{0.05} \right)^{-\xi} = -\mu + \sigma_{t+1} \times 1.69 \times \left(\frac{1}{5} \right)^{-0.25} = -\mu + \sigma_{t+1} \times 2.527.$$

In practice, we need to replace μ and σ_{t+1} by their estimates, which leads to the required result.

3. (a) The conditional portfolio variance is

$$\begin{aligned}
\sigma_{t+1}^2 &= w_1^2 \sigma_{11,t+1} + w_2^2 \sigma_{22,t+1} + 2w_1 w_2 \sigma_{12,t+1} \\
&= w_1^2 \left(0.02 + 0.03 R_{1t}^2 + 0.95 \sigma_{11t} \right) + w_2^2 \left(0.02 + 0.03 R_{2t}^2 + 0.95 \sigma_{22t} \right) \\
&\quad + 2 w_1 w_2 \left(0.01 + 0.03 R_{1t} R_{2t} + 0.95 \sigma_{12t} \right) \\
&= 0.02 \left(w_1^2 + w_2^2 + w_1 w_2 \right) + 0.03 (w_1^2 R_{1t}^2 + w_2^2 R_{2t}^2 + 2 w_1 w_2 R_{1t} R_{2t}) \\
&\quad + 0.95 (w_1^2 \sigma_{11t} + w_2^2 \sigma_{22t} + 2 w_1 w_2 \sigma_{12t}) \\
&= \omega + \alpha R_t^2 + \beta \sigma_t^2,
\end{aligned}$$

with $\omega = 0.02 \left(w_1^2 + w_2^2 + w_1 w_2 \right)$, $\alpha = 0.03$ and $\beta = 0.95$.

(b) Using σ_{t+1}, and the $N(0,1)$ random variables $\check{z}_{i,j}, i = 1, \ldots, MC; j = 1, 2$, we carry out the following for $i = 1, \ldots, MC$:

i. $\check{R}_{i,t+1} = \sigma_{t+1} \check{z}_{i,1}$;
ii. $\check{\sigma}_{i,t+2}^2 = \omega + \alpha \check{R}_{i,t+1}^2 + \beta \sigma_{t+1}^2$;
iii. $\check{R}_{i,t+2} = \check{\sigma}_{i,t+2} \check{z}_{i,2}$;
iv. $\check{R}_{i,t+1:t+2} = \check{R}_{i,t+1} + \check{R}_{i,t+2}$.

Note that ω follows from the portfolio weights, see (a), and $\alpha = 0.03$ and $\beta = 0.95$. With the resulting MC replications of $\check{R}_{i,t+1:t+2}$, we calculate

$$VaR_{t+1:t+2}^{0.01} = -Percentile(\left\{ \check{R}_{i,t+1:t+2}, i = 1, \ldots, MC \right\}, 0.01).$$

(c) The approximation is based on two assumptions:

i. $E_t(\sigma_{t+2}^2) = \sigma_{t+1}^2$, so that $\mathrm{Var}_t(R_{t+1:t+2}) = \sigma_{t+1}^2 + E_t(\sigma_{t+2}^2) = 2\sigma_{t+1}^2$;
ii. The distribution of $z_{t+1:t+2} = R_{t+1:t+2} / \sqrt{\mathrm{Var}_t(R_{t+1:t+2})}$ is standard normal.

The first assumption is valid when σ_{t+1}^2 is equal to the unconditional variance $\sigma^2 = \omega/(1 - \alpha - \beta)$. The approximation $2\sigma_{t+1}^2$ will over-estimate the risk when $\sigma_{t+1}^2 > \sigma^2$, and under-estimate it when $\sigma_{t+1}^2 < \sigma^2$ (because of mean-reversion).
The second assumption is never valid, even if $z_{t+j} \sim N(0,1)$. In practice $z_{t+1:t+2}$ has a higher kurtosis than the normal distribution, which implies that the approximation will under-estimate the risk. This second effect is stronger than the first (because the volatility mean-reversion is very weak), hence in general the approximation will under-estimate the true $VaR_{t+1:t+2}^{0.01}$.

4. (a) True: Backtesting expected shortfall is based on the exceedances: the observations where the return was less than minus the VaR. If we have a test period of T observations, and $p = 0.01$, this means that our backtest is based on only $0.01T$ observations (e.g., 4 years of data implies $T = 1000$, hence only 10 exceedances on average). So we need T to be very large to get a reasonable test sample. (For backtesting VaR, we use all T observations.)

(b) False: GJR-GARCH is a model for the leverage effect, i.e., asymmetry of the news impact function; this has nothing to do with skewness of a distribution.

(c) False: The realized variance is an ex post measure of risk, cannot be directly used to predict tomorrow's risk. It can be used to improve the GARCH model, by adding the realized variance as an explanatory variable.

(d) True: the Gaussian copula has zero tail dependence (regardless of ρ), whereas the Student's t copula has non-zero tail dependence (depending on the parameters d and ρ).

(e) False: the Vasicek model has a common factor that drives defaults. As the number of loans in the portfolio increases, the idiosyncratic credit risk is diversified away, but the common (systematic) risk does not vanish.