Yunchao Liu

Part 1: You own the 3 portfolios in the file “initial_portfolio.csv.” The risk free rate is in “rf.csv.” Daily prices of the stocks are in “DailyPrices.csv.” You bought these portfolios at the end of 2023. Model the returns of stocks using CAPM with SPY as the market. Use the data up to the end of 2023 for the regression. Your holding period on these portfolios is to the end of the price data. Use you the fitted models to attribute the realized risk and return for each portfolio and the total portfolio for the holding period. Split the attribution between the systematic and idiosyncratic components. You should calculate the idiosyncratic contribution for each stock, but present the total in your output. Discuss the results.

```
=== Total Portfolio Attribution ===
# 3x4 DataFrame
# ----------------------------------------------------------------------
#  Row | Value                        SPY          Alpha     Portfolio
# ----------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.05664199    0.20473092
#  2   | Return Attribution    0.25225478    -0.04284574    0.20473092
#  3   | Vol Attribution       0.00765925    0.00236786    0.00715334

=== A Portfolio Attribution ===
# 3x4 DataFrame
# ----------------------------------------------------------------------
#  Row | Value                        SPY          Alpha     Portfolio
# ----------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.12473119    0.13664173
#  2   | Return Attribution    0.25082955    -0.11203901    0.13664173
#  3   | Vol Attribution       0.00785466    0.00328149    0.00753042

=== B Portfolio Attribution ===
# 3x4 DataFrame
# ----------------------------------------------------------------------
#  Row | Value                        SPY          Alpha     Portfolio
# ----------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.05784708    0.20352583
#  2   | Return Attribution    0.24177638    -0.03564404    0.20352583
#  3   | Vol Attribution       0.00728329    0.00333397    0.00689811

=== C Portfolio Attribution ===
# 3x4 DataFrame
# ----------------------------------------------------------------------
#  Row | Value                        SPY          Alpha     Portfolio
# ----------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    0.01979954    0.28117245
#  2   | Return Attribution    0.26454517    0.02627267    0.28117245
#  3   | Vol Attribution       0.00783643    0.00295575    0.00797124
```

The return on the portfolio was around 5.66% lower than SPY. Most of the return (0.2523) was due to the market position (SPY). Negative alpha indicates an idiosyncratic loss, indicating that active contributions in addition to market exposure reduced returns. Subportfolio A had the worst return relative to SPY, with a significantly negative alpha (-0.1247). This indicates that sub-portfolio A securities significantly underperformed the market, even after taking into account the beta position. Sub-portfolio B largely reflects the entire portfolio. It captures most of the market return but still loses value through negative alpha, albeit less than A. Its volatility is slightly lower than A, indicating a more stable risk position. C is the most successful. It not only outperformed SPY with a positive total return of 28.12%, but also provided positive alpha, meaning that active stock selection or over/underweighting decisions added value that outperformed the beta position. Despite slightly higher volatility, this alpha return justifies the risk.

Part 2: Use your fitted CAPM results from Part 1, assume 0 alpha, and the expected return of the SPY is the average prior to the holding period. Assume the average risk free rate prior to the holding period is the expected risk free rate for the holding period. Create the optimal maximum Sharpe Ratio portfolio for each sub portfolio. Rerun the attribution from Part 1 using the new optimal portfolios. Discuss the results comparing back to Part 1. Given the fitted CAPM you have an expectation of the idiosyncratic risk contribution for each stock. How does the model compare to the realized values?

```
Portfolio A Optimal Allocation:
  Annualized Return: 20.13%
  Annualized Volatility: 13.80%
  Sharpe Ratio: 1.46

Portfolio B Optimal Allocation:
  Annualized Return: 20.06%
  Annualized Volatility: 13.56%
  Sharpe Ratio: 1.48

Portfolio C Optimal Allocation:
  Annualized Return: 20.22%
  Annualized Volatility: 13.68%
  Sharpe Ratio: 1.48
```

Total Portfolio Comparison:

Metric	Original Portfolio	Optimized Portfolio	Difference
Total Return	20.4700%	28.3900%	7.9200%
Systematic Return	24.9300%	26.4400%	1.5100%
Idiosyncratic Return	-4.4600%	1.9500%	6.4100%
Portfolio Beta	0.9500	1.0100	0.0600
Sharpe Ratio	-	1.476300	-

Comparison for Portfolio A:

Metric	Original Portfolio	Optimized Portfolio	Difference
Total Return	13.6600%	28.8600%	15.2000%
Systematic Return	25.2900%	26.4100%	1.1200%
Idiosyncratic Return	-11.6300%	2.4500%	14.0800%
Portfolio Beta	0.9700	1.0100	0.0400
Sharpe Ratio	-	1.463500	-

Comparison for Portfolio B:

Metric	Original Portfolio	Optimized Portfolio	Difference
Total Return	20.3500%	25.7900%	5.4400%
Systematic Return	24.0700%	26.3200%	2.2500%
Idiosyncratic Return	-3.7200%	-0.5300%	3.1900%
Portfolio Beta	0.9200	1.0100	0.0900
Sharpe Ratio	-	1.483600	-

Comparison for Portfolio C:

Metric	Original Portfolio	Optimized Portfolio	Difference
Total Return	28.1200%	30.5900%	2.4700%
Systematic Return	25.4300%	26.5900%	1.1600%
Idiosyncratic Return	2.6800%	4.0000%	1.3200%
Portfolio Beta	0.9700	1.0200	0.0500
Sharpe Ratio	-	1.482700	-

Part 3: Investigate the Normal Inverse Gaussian and the Skew Normal distributions. Explain how these distributions apply to finance, especially in relation to this class. you should look closely at the properties of each distribution and think about why they could be useful in Risk Management. The question is purposefully general to encourage critical thinking rather than a purely technical response. 1-2 pages

The Gaussian inverse normal (NIG) distribution is a member of the family of generalized hyperbolic distributions and is defined by four parameters: location (center), scale (dispersion), shape, and skewness. What makes the NIG particularly useful in finance is that it can model both fat tails and skewness of returns—two phenomena that are common in real-world financial data but are not captured by the standard bell curve of the normal distribution. In the real world, financial returns are neither obvious nor symmetric. Markets crash, economic recoveries are uneven, and strange events occur more often than you would expect if everything followed a normal distribution. This is where the NIG model excels. This gives greater weight to the possibility of extreme losses or gains (so-called "fat tails"), making it easier to model events such as financial crises or large stock market shocks. Asymmetry can also be dealt with well. So if a stock falls faster than it rises (which often happens), the NIG may reflect that more accurately.

One of the main advantages of the NIG method is its mathematical ease of use, which integrates well into advanced financial models. For example, it is often used in Lévy process models that account for asset price jumps, a significant improvement over standard Brownian motion. It is also used in derivatives pricing, particularly for illiquid assets or those subject to large price fluctuations. In the area of credit risk, the NIG method allows for modeling scenarios where default risk is unevenly distributed or losses may increase unexpectedly.

The NIG method is also useful for calculating value at risk (VaR) and expected loss (EL). In portfolios with many options or less liquid assets, standard deviation simply does not provide a complete picture. NIG-based risk measures provide a better picture of what could go wrong. It is also widely used in stochastic volatility models and for forecasting returns during turbulent and unpredictable market times.

Now let's compare this to the skewed normal (SN) distribution. This is a simpler extension of the normal distribution. It simply adds an extra parameter to skew the distribution to the left or right – useful when markets react more strongly to bad news than to good news, which is often the case. Although the SN is less flexible than the NIG, it is easy to use and quick to calculate, making it an indispensable tool for many everyday financial tasks. When the skewness parameter is zero, the SN simply reverts to the normal distribution and can be easily integrated into existing models. It is used to forecast returns, estimate credit default risk, and run Bayesian optimization models that incorporate conservative or pessimistic assumptions.

These two models, NIG and SN, proved extremely useful during the COVID-19-induced stock market crash of March 2020. The S&P 500 fell more than 30% in less than a month, far more than a normal distribution would have predicted. What followed was a slow and uneven recovery. Standard models couldn't accurately explain this. However, institutions using the NIG were better equipped: they were able to model the steep decline and distorted recovery more realistically. Their risk models were able to simulate these extreme events, allowing them to respond more quickly and effectively.

At the same time, biased normal models helped financial analysts capture fluctuations in investor sentiment. The avalanche of bad news (rising case numbers, lockdowns, economic crisis) caused markets to trend lower. The SN distribution allowed forecasters to account for this bias in their return estimates and stress tests.

Banks, and especially their credit risk departments, have begun to incorporate NIG and SN distributions into their default probability models. Think of small businesses, airlines, or hotels at the beginning of the pandemic: their revenues collapsed overnight, and their recovery was uncertain and slow. Standard symmetric models were not up to the task. However, NIG and SN distributions can better model these asymmetric risks.

In short, distributions like NIG and SN are essential if you want to understand and manage financial risk, especially when markets deviate from the norm. They are much more effective than the normal distribution at modeling extreme risks, asymmetries, and real-world behavior. Whether you're pricing derivatives, managing portfolios, or assessing credit risk, these models provide a more accurate and realistic picture of what could happen.

Part 4: Implement the Normal Inverse Gaussian and Skew Normal distributions (you can use implemented distributions in your stats package if exist). Using the pre-holding period data, create a risk model fitting each stock to the Normal, Generalized T, Normal Inverse Gaussian, and the Skew Normal choosing the best fit for each stock. Make the assumed return on each stock to be 0%. Report the best fit model for each stock and the parameters. Calculate the 1 day VaR and ES for each portfolio and the total portfolio using a Gaussian Copula and the fitted models. Do the same assuming a multivariate normal. Discuss the difference between the two approaches.

```
=== 1-Day VaR and ES (MVN vs Approx. Copula) ===
  Portfolio  VaR_MVN (5%)  ES_MVN (5%)  VaR_Copula (5%)  ES_Copula (5%)
0         A      0.012493     0.015962         0.013050        0.017111
1         B      0.011439     0.014544         0.011901        0.016149
2         C      0.012021     0.015514         0.012961        0.017151
3     Total      0.011769     0.014870         0.012417        0.016520
```

Copula-based risk measures are systematically higher. In all portfolios, VaR and especially ES are higher when calculated using the Gaussian copula approach. Linear correlation (MVN) underestimates the probabilities of joint extreme losses. Portfolio C appears to be the riskiest. Both the copula and MVN indicators show that Portfolio C has the highest 1-day ES (Copula ES = 1.72%). This means that Portfolio C has the most tail-dependent or volatile composition, possibly due to correlated stocks, high-beta stocks, or skewed composition. Portfolio B has the lowest risk indicators in both models. Even in the copula model, VaR and ES remain relatively low, suggesting a diversified or stable allocation. The total portfolio has a lower VaR/ES than any individual portfolio. This demonstrates that the diversification benefits persist even in the copula model. However, from a risk management perspective, the difference between MVN and Copula (combined ES: 1.49% versus 1.65%) remains significant.

	BestFit	Parameters
AAPL	GenNorm	(1.4992113715953974, 0, 0.014695556714566356)
NVDA	GenNorm	(1.1293079059994866, 0, 0.024972385622384324)
MSFT	GenNorm	(1.4320683392705071, 0, 0.017711989363310043)
AMZN	GenNorm	(1.292419741678591, 0, 0.020949499035933072)
META	GenNorm	(0.9967470882867913, 0, 0.016009309353142494)
GOOGL	GenNorm	(1.1863661273788813, 0, 0.01720077161863124)
AVGO	GenNorm	(1.107109748217272, 0, 0.016594696064111243)
TSLA	GenNorm	(1.387049198787663, 0, 0.03602418091115731)
GOOG	GenNorm	(1.221453989075414, 0, 0.017970966469120657)
BRK-B	GenNorm	(1.4084882520039237, 0, 0.009457713178621752)
JPM	GenNorm	(1.1276018596222461, 0, 0.010862812270947549)
LLY	GenNorm	(0.9491433474843236, 0, 0.011167181580656563)
V	GenNorm	(1.6051913210494368, 0, 0.012097767154269998)
XOM	GenNorm	(1.5060315104881745, 0, 0.01828089471936116)
UNH	GenNorm	(1.0735008939117512, 0, 0.010149274775364044)
MA	GenNorm	(1.2628054985963502, 0, 0.010587719501090373)
COST	GenNorm	(1.2055545783241512, 0, 0.011079264935161285)
PG	GenNorm	(1.2917151085849994, 0, 0.009427337299572881)
WMT	GenNorm	(1.2390974636144692, 0, 0.009042692779778521)
HD	GenNorm	(1.1859731105901092, 0, 0.012176403718466997)
NFLX	GenNorm	(1.0316857891797562, 0, 0.01691206997779695)
JNJ	GenNorm	(1.0417826180586411, 0, 0.007640908424385265)
ABBV	GenNorm	(1.0952084004963138, 0, 0.009752943158468462)
CRM	GenNorm	(1.099718588700762, 0, 0.015142737935547546)
BAC	GenNorm	(1.2114334771005564, 0, 0.015721129650622775)
ORCL	GenNorm	(0.8807512993411745, 0, 0.00982623330645517)
MRK	GenNorm	(1.5097593643609817, 0, 0.013886389686867255)
CVX	GenNorm	(1.2608532136878758, 0, 0.014053835696170539)
KO	GenNorm	(1.2826406068103655, 0, 0.008324721472792458)
CSCO	GenNorm	(1.0837185439351744, 0, 0.009758762688102378)
WFC	GenNorm	(1.2914475400731122, 0, 0.01744857141837457)

ACN	GenNorm	(1.3812746598157202, 0, 0.014786874213948406)
NOW	GenNorm	(1.2434648122135705, 0, 0.019309014625785447)
MCD	GenNorm	(1.478332639342592, 0, 0.010144612934366798)
PEP	GenNorm	(1.2430690743648107, 0, 0.00905229658796294)
IBM	GenNorm	(1.2409940537414255, 0, 0.00945409902547802)
DIS	GenNorm	(1.2377061014199597, 0, 0.01575029840678447)
TMO	GenNorm	(1.336237187552771, 0, 0.01485743291779756)
LIN	GenNorm	(1.1065668246574658, 0, 0.010213041973584143)
ABT	GenNorm	(1.3571831723127596, 0, 0.012997122433684803)
AMD	GenNorm	(1.1707846788571343, 0, 0.026442051909039165)
ADBE	GenNorm	(1.2815565538182612, 0, 0.020088661987634662)
PM	GenNorm	(1.4881389040170805, 0, 0.011976480505092445)
ISRG	GenNorm	(1.1692913393553659, 0, 0.016042597149610323)
GE	GenNorm	(1.457002330946111, 0, 0.017471317509183095)
GS	GenNorm	(1.3248564721514529, 0, 0.015628886090369148)
INTU	GenNorm	(1.2436449606856097, 0, 0.01780826015841748)
CAT	GenNorm	(1.2205939900517566, 0, 0.016640672556286946)
QCOM	GenNorm	(1.28295479313674, 0, 0.019635744690017047)
TXN	GenNorm	(1.604217136162101, 0, 0.018537458957011484)
VZ	GenNorm	(0.999728031025005, 0, 0.009728560891713686)
AXP	GenNorm	(1.1512284721995687, 0, 0.01411794624474901)
T	GenNorm	(0.9941939720360091, 0, 0.01114720290793745)
BKNG	GenNorm	(1.5086127033393386, 0, 0.018462078667605332)
SPGI	GenNorm	(1.200445132483253, 0, 0.012530688958509877)
MS	GenNorm	(1.178487592079398, 0, 0.014566300528911367)
RTX	GenNorm	(0.9085831084912563, 0, 0.008450694982345389)
PLTR	GenNorm	(0.9175395370424646, 0, 0.026050597848529)
PFE	GenNorm	(1.2243797513539192, 0, 0.013638398217172382)
BLK	GenNorm	(1.4584897188950718, 0, 0.015798680418226782)
DHR	GenNorm	(1.313715294866354, 0, 0.015421363261090101)
NEE	GenNorm	(1.0193912073547835, 0, 0.012389275628415805)
HON	GenNorm	(1.3599334556194667, 0, 0.01216441124777376)
CMCSA	GenNorm	(1.1307237158027421, 0, 0.01229683230022319)
PGR	GenNorm	(0.9037019917160931, 0, 0.010133627298517326)
LOW	GenNorm	(1.117211910325571, 0, 0.01246926497326609)
AMGN	GenNorm	(1.3488893503679307, 0, 0.01373063505303486)
UNP	GenNorm	(1.0416601089316826, 0, 0.010721373183964138)
TJX	GenNorm	(1.5873146863904362, 0, 0.012328367553501744)
AMAT	GenNorm	(1.3766639351348253, 0, 0.02309592962481167)
UBER	SkewNormal	(0.20451426135530312, 0, 0.02304565300345477)
C	GenNorm	(1.1462071168243693, 0, 0.013838063137472746)
BSX	GenNorm	(1.0586086116032676, 0, 0.009644960719979914)
ETN	GenNorm	(1.2010223585544089, 0, 0.015325268764271123)
COP	GenNorm	(1.374054166697991, 0, 0.019138895230821004)

```
BA       GenNorm  (1.1901034673923199, 0, 0.015505932532510185)
BX       GenNorm  (1.3679187730467568, 0, 0.023459933498004855)
SYK      GenNorm  (0.9224647444868863, 0, 0.008770609516560493)
PANW     GenNorm  (1.0270219784870278, 0, 0.01735456094734085)
ADP      GenNorm  (0.9853041466935804, 0, 0.008761345029646816)
FI       GenNorm  (1.0610871875018357, 0, 0.009994454752279338)
ANET     GenNorm  (0.8366271074777182, 0, 0.013431305916266924)
GILD     GenNorm  (1.4660486728394906, 0, 0.014558931119064837)
BMY      GenNorm  (1.201728982112971, 0, 0.011157732204996051)
SCHW     GenNorm  (0.9143736455952021, 0, 0.01546163360039304)
TMUS     GenNorm  (1.3392376814118716, 0, 0.012315963498079712)
DE       GenNorm  (1.3374999565345642, 0, 0.017721422391752437)
ADI      GenNorm  (1.3368403152529673, 0, 0.017085025196080447)
VRTX     GenNorm  (1.0398545792384544, 0, 0.011186511914125037)
SBUX     GenNorm  (1.0729606456028717, 0, 0.010645481917577693)
MMC      GenNorm  (1.2574492139750324, 0, 0.010251326749755242)
MDT      GenNorm  (1.2441298401959044, 0, 0.012914651661944495)
CB       GenNorm  (1.31389647137439, 0, 0.013050157760401788)
LMT      GenNorm  (1.0144473289345113, 0, 0.007778725463773959)
KKR      GenNorm  (1.459413794467985, 0, 0.022797102533004 66)
MU       GenNorm  (1.1991303772534165, 0, 0.021052392752855 94)
PLD      GenNorm  (1.4293807624380994, 0, 0.01851232000099712)
LRCX     GenNorm  (1.2334297517898714, 0, 0.02194804858936275)
EQIX     GenNorm  (1.279561517949322, 0, 0.015164540235462275)
```

For each stock (except SPY), four different probability distributions (normal, skewed, generalized normal, and inverse normal Gaussian) are fitted by estimating their parameters and calculating the log likelihood of the observed data for each fitted model. The log likelihood indicates how well a distribution explains the observed returns, with higher values indicating better fits. The script then selects the best-fitting model for each stock based on the highest log likelihood and records the distribution name and associated parameters.

Part 5: Using your best fit risk model, calculate a risk parity portfolio for each sub portfolio using ES as the risk metric. Rerun the attribution from Part 1 using the new optimal portfolios and the previously fit CAPM beta. Discuss the results comparing back to Part 1 and Part 2.

```
=== Total Portfolio Attribution ===
# 3x4 DataFrame
# ----------------------------------------------------------------------
#  Row | Value                        SPY          Alpha      Portfolio
# ----------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.02477541    0.23659750
#  2   | Return Attribution    0.25111748    -0.00632898    0.23659750
#  3   | Vol Attribution       0.00748455    0.00327751    0.00731296
```

```
=== A Portfolio Attribution ===
# 3x4 DataFrame
# ---------------------------------------------------------------------
#  Row | Value                          SPY         Alpha      Portfolio
# ---------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.08517817    0.17619474
#  2   | Return Attribution    0.24225688    -0.06285888    0.17619474
#  3   | Vol Attribution       0.00745167     0.00376273    0.00729006

=== B Portfolio Attribution ===
# 3x4 DataFrame
# ---------------------------------------------------------------------
#  Row | Value                          SPY         Alpha      Portfolio
# ---------------------------------------------------------------------
#  1   | TotalReturn           0.26137291    -0.05350747    0.20786544
#  2   | Return Attribution    0.25182513    -0.03994630    0.20786544
#  3   | Vol Attribution       0.00757204     0.00358583    0.00724754

=== C Portfolio Attribution ===
# 3x4 DataFrame
# ---------------------------------------------------------------------
#  Row | Value                          SPY         Alpha      Portfolio
# ---------------------------------------------------------------------
#  1   | TotalReturn           0.26137291     0.02314401    0.28451693
#  2   | Return Attribution    0.25456673     0.04336311    0.28451693
#  3   | Vol Attribution       0.00743885     0.00516418    0.00856001
```

This code implements portfolio performance analysis by integrating a CAPM-based attribution into a risk parity weighting strategy using expected shortfall (EPD). First, historical price data, portfolio allocations, and the risk-free rate are loaded. Then, daily returns and excess returns (returns less the risk-free rate) prior to 2024 are calculated. For each stock, a T-distribution is fitted to the excess returns to estimate the 5% expected shortfall, which serves as a measure of downside risk. These DE values allow the calculation of risk parity weights: assets with lower downside risk are assigned higher weights. The initial positions of the CSV portfolio are deleted and replaced with these weights, which are normalized within each sub-portfolio (A, B, and C). Using these new weights, the code fits a CAPM model for each asset to estimate alpha and beta. It then divides the portfolio's returns and volatility into systematic (market) and idiosyncratic (alpha/residual) components. The result shows the extent to which portfolio performance is determined by market exposure versus individual asset selection, both for the overall portfolio and for each sub-portfolio.

A comparison of the three portfolio allocation approaches (regular, CAPM-optimized, and ES-based risk parity) highlights clear trade-offs between improved returns and risk control. The CAPM-optimized portfolio consistently outperforms the others, delivering the highest total returns and converting negative alpha into significant positive idiosyncratic contributions, especially for underperforming portfolios like Portfolio A. This reflects the benefit of aligning portfolio weights with systematic exposures while neutralizing alpha assumptions. In contrast, the ES-based risk parity approach focuses on controlling downside risk by equalizing tail risk contributions. This results in better volatility attribution and a more balanced exposure, but slightly lower returns than its CAPM-driven counterpart. The regular (non-optimized) portfolios exhibit significantly worse performance due to high negative alpha, indicating insufficient initial alignment with market factors. In general, CAPM optimization is ideal for maximizing returns in a linear risk model, while ES-based risk parity offers a more conservative and balanced risk allocation strategy.