1) **Backtests (final):**

All are satisfying criteria for pvalues

- **For 10 day step size from 2000 to 2025 for two stocks :**

📊 Backtest Results (VaR Level: 5%)

VaR violations: 30 / 521 (5.76%)
Kupiec Test p-value: 0.4576
Christoffersen p-value: 0.5255
Acerbi–Szekely ES stat: 2.1215
Acerbi–Szekely crit value: 2.45873320537428
Acerbi–Szekely ES p-value: 0.2386
Mean Absolute ES Error: 0.5819(0.570789 with garch method, error is almost same but christoffersen value is 0.06 whereas our model has 0.52 so our model is better than garch model(where there is no gpd and copula))

- **For one day step size from 2020 to 2025 for two stocks :**

As calculating every day from 2000 to 2025 takes more time, 2020 to 2025 period considered which gives 260 observations

📊 Backtest Results (VaR Level: 5%)

VaR violations: 12 / 261 (4.60%)
Kupiec Test p-value: 0.7563
Christoffersen p-value: 0.5683
Acerbi–Szekely ES stat: 2.0940
Acerbi–Szekely crit value: 2.685823754789272
Acerbi–Szekely ES p-value: 0.3343
Mean Absolute ES Error: 1.0569   (**1.257199** for garch method without gpd ,copula   so able to reduce error with our model)

Here 10 day step size has less es error but we should not consider it to compare with one day because 10 day es is smoother(considers every 10th day) and the data period is longer for 10 day so taking average gives less error than one day . and if we consider remaining values in every value one day step size has better results than 10 day step size.and also one day is giving better es error than baseline so one day step size is giving better results than 10 day step size .

- **For one day stepsize from 2020 to 2025 for 4 stocks**

  📊 Backtest Results (VaR Level: 5%)

  VaR violations: 12 / 261 (4.60%)
  Kupiec Test p-value: 0.7563
  Christoffersen p-value: 0.1046
  Acerbi–Szekely ES stat: 1.9987
  Acerbi–Szekely crit value: 2.685823754789272
  Acerbi–Szekely ES p-value: 0.4345
  Mean Absolute ES Error: 0.5819(0.627130)

- **For 10 day stepsize from 2000 to 2025 for 4 stocks**
  📊 Backtest Results (VaR Level: 5%)

  VaR violations: 26 / 521 (4.99%)
  Kupiec Test p-value: 0.9920
  Christoffersen p-value: 0.5462
  Acerbi–Szekely ES stat: 1.9978
  Acerbi–Szekely crit value: 2.45873320537428
  Acerbi–Szekely ES p-value: 0.5335
  Mean Absolute ES Error: 0.5220(0.594655)

  As far as 4 stocks are considered 10 day step size giving results better than one day step size

2)   Dependence modelling

i)For two assets(2000-2015)
I have done copula selection using aic criterion which checks for overfitting etc   . with this criteria I got t copula as the best . but if we prefer only gof p value bb7 is better than t copula .we want our model to fit well on unseen data, as gof p value tells us how well it fits on seen data so i thought it is better to consider aic criteria

Global GoF p-value: 0.1789 for student copula
Global GoF p-value: 0.4200 for bb7

With t copula the lambda estimates are
empirical_lambda_L:0.3299145299145299 , empirical_lambda_U:0.2674616695059625
sim_lambda_L:0.31918505942275044,   sim_lambda_U: 0.29604130808950085

The 95 percent confidence intervals lambdas are :
95% CI for λ_L: 0.2924344283167813, 0.3683426422667849
95% CI for λ_U: 0.23315930892031927, 0.30313588850174217

The lambda estimates lie in 95 percent confidence intervals . these are calculated considering t copula for the data from 2000 to 2015

ii)For four assets(2000-2015) :

Fitted vine copula and through aic criteria it selected student copula for 5 pairs and gaussian for one pair (vine copula selects copula for pairs seperately and then models)

Cramér–von Mises statistic: 0.000068
Rosenblatt GoF p-value (d=4): 0.1253

3) for the data from 2000 to 2015 ,During backtesting there are some scenarios where no tests are passing for gpd fit satisfying ks and ad p value greater than zero and no of exceedences greater than 100 then i did gpd fit to only those that passes gpd tests and for remaining data i have used kde. If left tail passes i did gpd to left tail but for remaining data i have used only kde for center and right tail

i)For two stocks

Marginal CDFs:

	Stock	Tail	Threshold (u)	Shape	Loc	Scale	KS Statistic
0	INFY.NS	upper	0.534919	0.040481	0	0.619985	0.034371
1	INFY.NS	lower	-0.708590	0.189417	0	0.508399	0.042487
2	SBIN.NS	upper	0.905630	0.003990	0	0.597632	0.017256
3	SBIN.NS	lower	-0.593254	0.016617	0	0.618967	0.034552

	KS p-value	adpvale	AD Statistic	Num Exceedances
0	0.224678	0.293333	1.167035	915
1	0.138279	0.076667	2.249612	732
2	0.993749	0.893333	0.324557	586
3	0.219601	0.296667	1.207604	915

Cumulative distribution by combining kde and gpd :

Pit values :

PIT Histogram for INFY.NS

PIT Histogram for SBIN.NS

= 0.1769
= 0.0610
re uniform
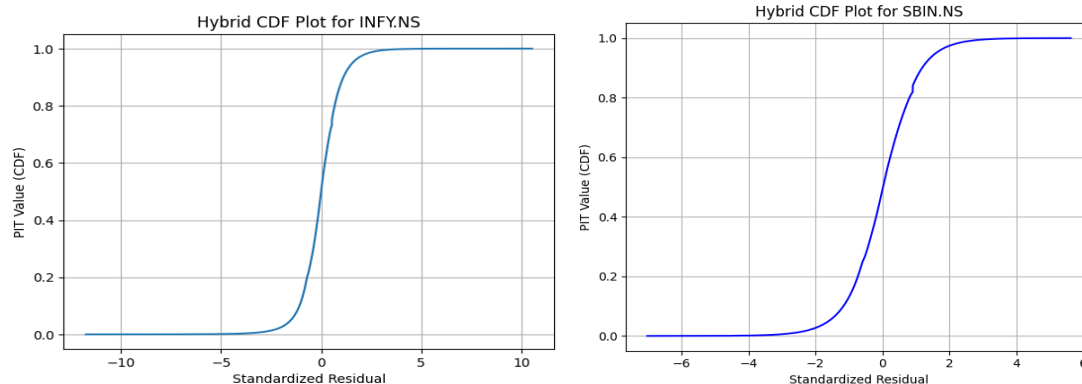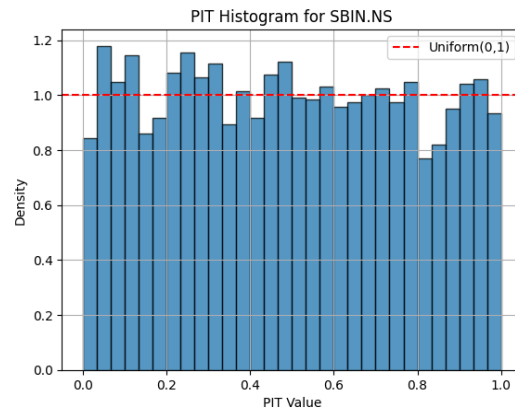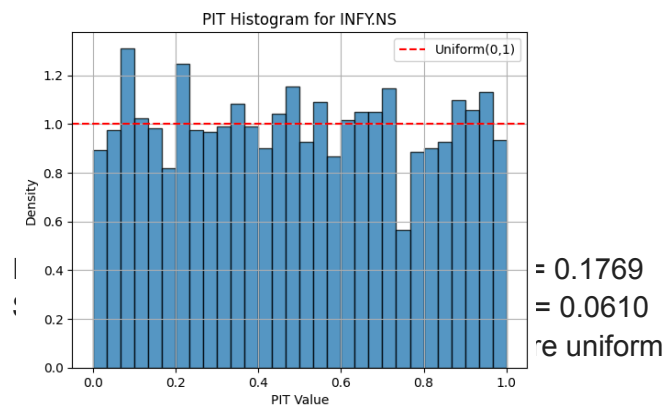
ii)For four stocks

Marginal CDFs:

	Stock	Tail	Threshold (u)	Shape	Loc	Scale	KS Statistic \
0	HUL.NS	upper	0.580636	0.007255	0	0.667998	0.035475
1	HUL.NS	lower	-0.564554	0.027587	0	0.592210	0.025505
2	INFY.NS	upper	0.535056	0.040594	0	0.619877	0.034666
3	INFY.NS	lower	-0.708082	0.188898	0	0.508943	0.042955
4	SBIN.NS	upper	0.905623	0.003990	0	0.597632	0.017248
5	SBIN.NS	lower	-0.593261	0.016550	0	0.619001	0.034531
6	TATAMOTORS	upper	1.950723	0.043737	0	0.482298	0.055984
7	TATAMOTORS	lower	-1.430466	0.023440	0	0.536646	0.038026

	KS p-value	adpvale	AD Statistic	Num Exceedances
0	0.195040	0.170000	1.508961	915
1	0.582072	0.423333	0.845823	915
2	0.216445	0.306667	1.201950	915
3	0.130374	0.060000	2.270573	732
4	0.993784	0.933333	0.323886	586
5	0.220163	0.240000	1.206016	915
6	0.861460	0.380000	0.977296	110
7	0.837297	0.653333	0.546964	257

Pit values :

HINDUNILVR.NS: KS statistic = 0.0179, p-value = 0.1874
INFY.NS: KS statistic = 0.0181, p-value = 0.1808

SBIN.NS: KS statistic = 0.0218, p-value = 0.0610
TATAMOTORS.NS: KS statistic = 0.0144, p-value = 0.4333
they are uniform and not clustered near 0 or 1

To do :
- For var lower tail dependence is important so i wanted to run once with clayton copula and compare results with student copula
- The way of selecting copula based on aic criteria if it is wrong then please let me know sir i will work on it
- Ks values for pit values are just greater than 0.05 for some stocks so i will try to improve it
- Historical method is also   giving results same as our model so i wanted to check once

Colab link for backtesting : ∞ Untitled21.ipynb

Colab link for finding gof values,thresholds for data from 2000 to 2015 :
 ∞ garch-evt-copula.ipynb