Benchmarking Low Default Portfolios to Third Party Ratings

Distance-Based Tendency Testing

Andrija Djurovic

www.linkedin.com/in/andrija-djurovic

Low Default Portfolios Benchmarks

- Validating low-default portfolios (LDP) presents significant challenges to practitioners, as most statistical tests applicable to high default portfolios do not have equivalents for low default portfolios.
- Therefore, practitioners often rely on a benchmarking approach to validate certain types of LDP where external ratings are available.
- Benchmarking helps assess how well the rating system distinguishes between low- and high-risk counterparties and whether the average risk assessment significantly differs from an independent observer's.
- It also evaluates whether deviations in ratings are unreasonably large.
- The drawbacks of benchmarking include that external ratings are typically available for only part of the portfolio, which may not represent the entire portfolio. Unlike default prediction, deviations from external ratings are not necessarily incorrect. Therefore, considering factors such as differences in rating and default definitions, careful analysis and detailed interpretation of benchmarking results are essential.
- A fundamental assumption of benchmarking is confidence in the correctness of external ratings. Accordingly, selecting the relevant type of external rating requires great care and an understanding of potential discrepancies.
- This presentation focuses on distance-based tendency testing.
- More on LDP benchmarking can be found in the following reference: Franz, C., & Lawrenz, C. (2007). Benchmarking low-default portfolios to third-party ratings. Journal of Credit Risk, 3(2), 87–100.

Distance-Based Tendency Testing

Tendency is one of the measures calculated based on the distances between internal and external ratings.

It evaluates the systematic bias in internal ratings compared to external ratings by categorizing deviations as optimistic, neutral, or pessimistic.

A prerequisite for testing tendency and other distance-based benchmarking procedures is the availability of a mapping process that translates internal ratings to external ratings.

The basis for the tendency calculation can be expressed as:

$$p_{-1} = \frac{N_{-1}}{N}, \quad p_0 = \frac{N_0}{N}, \quad p_{+1} = \frac{N_{+1}}{N}$$

where:

- N_{-1} is the count of debtors where the internal rating is worse than the external rating;
- N_0 is the count where internal and external ratings are equal;
- N_{+1} is the count where the internal rating is better than the external rating;
- N is the total number of debtors.

Distance-Based Tendency Testing cont.

Simultaneous confidence intervals for these proportions at a selected level α can be estimated using multinomial distribution approaches, such as Goodman's method:

$$[p_i^l, p_i^u] = \left[\frac{B_i - \sqrt{B_i^2 - 4AN_i^2/N}}{2A}, \frac{B_i + \sqrt{B_i^2 - 4AN_i^2/N}}{2A} \right]$$

where:

- A is defined as $\chi^2(1 - \frac{\alpha}{3}) + N$;
- B is defined as $\chi^2(1 - \frac{\alpha}{3}) + 2N_i$;
- with χ^2 is the quantile of the chi-square distribution for confidence level α with one degree of freedom.

Simulation Study

- The following slides provide `R` and `Python` code for testing tendency.
- The simulated dataset consists of internal and external ratings, the only inputs needed for this calculation.
- The dataset is available here.
- In addition to the proposed simultaneous confidence intervals for the multinomial proportions, practitioners should be aware that similar testing can be conducted using a test of two proportions, directly comparing p_{-1} and p_{+1} while disregarding the overlapping samples.

R Code

```r
#utils function import (mltnp.ci)
source("https://raw.githubusercontent.com/andrija-djurovic/adsfcr/refs/heads/main/ldp/utils_benchmarking_ldp.R")

#data import
url <- "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/refs/heads/main/ldp/ldp_benchmarking.csv"
db <- read.csv(file = url,
               header = TRUE)

#tendency indicator
tendency <- ifelse(db$internal > db$external, 1,
            ifelse(db$internal < db$external, -1, 0))

#frequency table
table(tendency)
```

```
## tendency
##  -1   0   1
## 349 173 478
```

```r
#tendency confidence intervals
mltnp.ci(tendency = tendency,
         cl = 0.95)
```

```
##       est     lower     upper
## -1 0.349 0.3138684 0.3858525
## 0  0.173 0.1462494 0.2034774
## 1  0.478 0.4404176 0.5158332
```

Python Code

```
import pandas as pd
import numpy as np
from scipy import stats
import requests

#utils function import (mltnp.ci)
url = "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/refs/heads/main/ldp/utils_benchmarking_ldp.py"
r = requests.get(url)
exec(r.text)

#data import
url = "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/refs/heads/main/ldp/ldp_benchmarking.csv"
db = pd.read_csv(filepath_or_buffer = url)

#tendency indicator
tendency = np.where(db["internal"] > db["external"], 1,
           np.where(db["internal"] < db["external"], -1, 0))

#frequency table
np.unique(ar = tendency,
          return_counts = True)
```

```
## (array([-1,  0,  1]), array([349, 173, 478], dtype=int64))
```

```
#tendency confidence intervals
mltnp_ci(tendency = tendency,
         cl = 0.95)
```

```
##       est     lower     upper
## -1  0.349  0.313868  0.385853
##  0  0.173  0.146249  0.203477
##  1  0.478  0.440418  0.515833
```