

CONFIDENCE INTERVALS FOR STRESS TEST PREDICTIONS

YAACOV KOPELIOVICH AND KEVIN SHEA

ABSTRACT. We obtain a formula that calculates confidence intervals for stress testing predictions. We show a numerical example comparing the result of our formula with a hypothetical scenario bootstrap methodology.

1. INTRODUCTION

Stress testing was introduced to the financial industry following the global financial crisis of 2008 to 2009. Recall that in 2008, the acceptable risk paradigm of value at risk (VAR) failed to produce adequate warnings to the financial calamity that occurred for many banks when the subprime crisis happened. The crisis impacted all layers of the society and forced the US government and the Federal Reserve to spend hundreds of billions of dollars to stabilize the balance sheets of the largest banks and insurance companies. As a result, the Federal Reserve introduced a new paradigm of stress testing that was institutionalized in the Dodd Frank Wall Street Reform and Consumer Protection Act.

On an annual basis, the Federal Reserve publishes different economic scenarios (a base scenario and an adverse scenario) in which are incorporated different shocks to the financial system. These shocks include changes to market factors like the Russell 1000 and High Yield Spreads, as well as economic factors like GDP and Unemployment. Banks run projections on their portfolios and return the results to the Federal Reserve. The results determine if the bank needs to take some capital actions or if their financial position is sound. The success of this risk framework was clearly seen in the Covid pandemic crisis when US banks had a stellar balance sheets and avoided the many issues they experienced during the 2008 subprime crisis.

A natural question relates to the "confidence intervals" of the projections supplied to the Federal Reserve by the banks. The reason for this, is that while a bank might pass a stress test with a base scenario as well as an adverse scenario, investors may want to know the probability, as determined through confidence intervals, of an existential threat. Put another way, while two banks might pass a stress test, under the adverse scenario, one bank might have a 0.01% chance of its book equity being fully depleted while another bank might have a 30% chance of having its book equity fully depleted. We think investors will want to know not just whether a bank passes a stress test but also the probability of an existential threat such as having its book equity being fully depleted.

While confidence intervals can be achieved through a bootstrapping method it may still be convenient to have an analytic formula at hand to calculate the confidence intervals. Such a formula can assist the auditors and the analysts to quickly validate the scenarios produced by the banks and can also serve as a confirming methodology if it can be implemented easily and efficiently. In this short note we suggest such a formula. Surprisingly we have not found this formula in any financial

literature or risk textbook. Due to its simplicity, and the simplicity of the underlying assumptions, we present it hoping that it will be of use to financial auditors and risk managers. The structure of this note is as follows:

In section 1 we present the formula and the underlying assumptions. In section 2 we will use the bootstrapping method to obtain confidence intervals on a benchmark test portfolio and compare the results with the formula we found. We summarize our findings and conclusions in section 3 and suggest directions for future research.

2. CONFIDENCE INTERVALS FOR STRESS TESTING SCENARIOS

We calculate the confidence interval for stress testing. Let F_1, \dots, F_n be the underlying factors and Y is the dependent variable. We assume that the factors F_1, \dots, F_n and Y follow multivariable normal distribution. By the usual regression assumptions we can write:

$$(1) \quad Y = \sum_{i=1}^n \beta_i F_i + \beta_0 + \epsilon$$

The financial shocks for F_1, \dots, F_n are known as they are provided by the Federal Reserve and can be written as: $\Delta F_1, \dots, \Delta F_n$. Using the regression assumptions we find the predicted change for Y is:

$$(2) \quad \Delta Y = \sum_{i=1}^n \beta_i \Delta F_i$$

As stated above we want to learn more about the distribution of ΔY . As the number of points utilized for the regressions is usually large (banks and the Federal Reserve often work with time series that are daily in nature) we assume that ΔY follows the normal distribution. Furthermore, as the shocks of $\Delta F_1, \dots, \Delta F_n$ are part of a scenario that we are really interested in, it is important to assume the distribution of Y is **conditional** on the shocks $\Delta F_1, \dots, \Delta F_n$. We assume that ΔY is normally distributed and hence to find the distribution of ΔY it suffices to find its expected value and its standard deviation **conditional** on $\Delta F_1, \dots, \Delta F_n$. The expected value is given as:

$$(3) \quad E(\Delta Y_{|\Delta F_1, \dots, \Delta F_n}) = \sum_{i=1}^n \beta_i \Delta F_i$$

To find the standard deviation of ΔY we take the following approach:

Lemma 1. Assume H_1, \dots, H_n are random variables with a covariance matrix C_H then the variance of $\sum_{i=1}^n w_i H_i$ is :

$$(4) \quad \sigma^2 \left(\sum_{i=1}^n w_i H_i \right) = W C_H W^t$$

$W = (w_1, \dots, w_n)$ is a $n \times 1$ vector and W^t is its transpose.

The proof of this lemma is straightforward by expanding the RHS using the properties of the Variance:

- (1) $\sigma^2(X_1 + X_2) = \sigma^2(X_1) + 2Cov(X_1, X_2) + \sigma^2(X_2)$
- (2) $\sigma(cX) = c\sigma(X)$

We apply this lemma to our situation. As $\Delta F_1, \dots, \Delta F_n$ are constants we have the following:

$$(5) \quad \sigma^2(\Delta Y_{|\Delta F_1, \dots, \Delta F_n}) = \Delta F C_\beta \Delta F_n^t$$

Here:

- (1) C_β = Covariance of the β coefficients

$$(2) \Delta F = (\Delta F_1, \dots, \Delta F_n)$$

From that we have the standard deviation of ΔY is:

$$(6) \quad \sigma(\Delta Y_{|\Delta F_1, \dots, \Delta F_n}) = \sqrt{\Delta F C'_\beta \Delta F_n^t}$$

We found the expected values and standard deviation of ΔY . We can use the formula to find the confidence level of α . i.e we are looking for a number y_0 such that $P(\Delta Y \leq y_0) = \alpha$, and $P(\Delta Y \leq y_0)$, is the probability of the random variable ΔY to be less or equal than y_0 . We have:

$$(7) \quad y_0 = E(\Delta Y) + Z_\alpha \sigma_Y$$

or

$$(8) \quad P(Y \leq E(\Delta Y) + Z_\alpha \sigma_Y) = \alpha$$

and Z_α is the usual Z score function.

The last formula is the main result of this note. It produces confidence intervals for stress test scenarios. As we indicated in the introduction despite the simplicity of this formula we haven't been able to find it in standard statistics books.

The formula above depends on the knowledge of C_β . (the covariance matrix of the regression coefficients.) The formula for C_β is complicated under the most general assumptions, however there is a case in which it simplifies and which will suffice for our needs. Assume conditionality of Y on the factors F_1, \dots, F_n then we have the following formula for C_β :

Lemma 2. *Assume that Y is distributed normally conditional on the factors F_1, \dots, F_n we have that:*

$$(9) \quad C_\beta = C_F^{-1} \frac{\sigma_\epsilon^2}{n}$$

and:

- (1) C_F = Covariance of factors
- (2) σ_ϵ = regression standard error
- (3) n = number of regression observations.

Combining the lemma with the discussion above demonstrates how to derive confidence intervals for shocks. In the next section we are going to see an example of this.

3. EXAMPLE

In this section we produce an example an example of a confidence interval for a stress test using the results of the last section.

Consider a hypothetical situation where the market and economic factors are the changes in the **Russell 1000**'s index and the changes in the Merrill Lynch **High Yield Spread** index ¹. We examine the results of this shock on the equity value of banks. As a proxy for bank equity, we use the index IAT, a regional bank ETF. We model the impact of a 1% change in the Russell 1000 and a 1% change in the Merrill Lynch High Yield Spread Index on IAT. We downloaded the time series of the Russell 1000, the Merrill Lynch High Yield Spread, and the IAT from Factset and used the R programming language to perform some analysis. The annual covariance matrix return for the market and economic factors was:

$$(10) \quad C = \begin{pmatrix} 19.91 & -2.83 \\ -2.83 & 0.726 \end{pmatrix}$$

In this matrix the first row corresponds to the **Russell 1000** factor and the second corresponds to **High Yield Spreads** The number of observations for 10 years is

¹The source of the data for various indices is the Federal Reserve St.Louis FRED.

185 and the annualized standard error is: 5.038 If C_β denotes the covariance of the regression coefficients β we have that:

$$(11) \quad C_\beta = C^{-1} \frac{\sigma_\epsilon^2}{n}$$

Substituting into this formula we have that the covariance matrix for the beta coefficients are:

$$(12) \quad C_\beta = \begin{pmatrix} 0.0153 & 0.0597 \\ 0.0597 & 0.4192 \end{pmatrix}$$

According to our hypothetical scenario of a 1% change in the Russell 1000 and a 1% change in the Merrill Lynch High Yield Spread Index we calculate the standard deviation of the projected distribution in IAT according to the formula (6) :

$$(13) \quad 0.74 = \sqrt{u C_\beta u^t}$$

Where $u = (0.01, 0.01)$.

We describe the bootstrapping algorithm that reconstructs our formula:

- (1) Simulate the residual as a normal variable with σ_ϵ . Call the matrix of this simulation *res*
- (2) Simulate the shock variable using
$$S = \alpha + \beta_1 F_1 + \beta_2 F_2 + res$$
- (3) For each simulation path of stock variable run a regression for the factors F_1, F_2 . to obtain simulated β_{1i}, β_{2i}
- (4) For these $\beta_{ij}, j = 1, 2$ create a simulated shock vector by $\beta_{1i} \Delta F_1 + \beta_{2i} \Delta F_2$
- (5) Calculate the standard deviation.

Results that we obtain for our example are:

- **0.75**
- **0.74**

These results are very close as expected.

4. CONCLUSION

We have produced a formula to calculate confidence intervals for stress scenarios. We saw that if we use a bootstrapping approach our formula produces a good approximation to the simulated result. In our opinion our formula can be viewed a benchmark if a bank wants to produce confidence intervals so the regulators, and investors for that matter, will be able to examine the statistical properties of the shock and not just the expected value. Our assumptions rely heavily on conditional probabilities. We intend to produce formulas omitting conditional probability assumptions in subsequent notes.

REFERENCES

- [HT] Hogg and Tanis PROBABILITY AND STATISTICAL INFERENCE Fifth Edition, Prentice Hall, 1994
- [FR] Federal Reserve Stress Test scenarios <https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20210212a1.pdf>

DEPARTMENT OF FINANCE, SCHOOL OF BUSINESS, UNIVERSITY OF CONNECTICUT, STORRS

DISCIPLINED ALPHA, ONE MARINA DRIVE STE. 1490, BOSTON MA 02210

E-mail address: yaacov.kopeliovich@uconn.edu, kevin.shea@disciplinedalpha.com