SAS/ETS® SEVERITY Procedure for Loss Modeling

Model the Entire Distribution of Continuous-Valued Events

Overview

Modeling the probability distribution of the magnitude of losses enables insurance and financial companies to plan for future losses and to meet minimum capital requirements under regulations such as Basel III and Solvency II. The SEVERITY procedure in SAS/ETS enables you to model these types of loss distributions. You can use PROC SEVERITY to fit any continuous distribution to the historical loss data by minimizing an objective function of your choice while accounting for censoring, truncation, and regression effects. The SEVERITY procedure produces several output tables and diagnostic plots to assess the adequacy of a given distribution model and to choose the best model to describe past losses.

You can use the SEVERITY procedure to perform the following types of analyses:

- Fit 10 predefined distributions: Burr, exponential, gamma, generalized Pareto, inverse Gaussian (Wald), lognormal, Pareto, scaled-Tweedie, Tweedie, and Weibull.
- Programmatically define and fit your own distributions, such as mixture distributions, and specify your own objective function to estimate parameters.
- Model regression effects on the scale or log-transformed scale parameter of a distribution.
- Account for censoring and truncation effects including interval-censoring (grouped data).
- Fit several distribution models at one time and choose the best model by comparing them across seven different statistics of fit.
- Conduct nonparametric analysis by estimating the empirical distribution function (EDF) and its standard errors. If you specify censoring or truncation, PROC SEVERITY uses either Kaplan-Meier’s or Turnbull’s algorithm.
- Compare empirical and parametric estimates visually with PDF, CDF, P-P, and Q-Q plots.
- Fit models faster on multicore CPUs with built-in multithreading.

In addition, you can perform useful tasks that are related to loss distribution modeling, including the following:

- Use the INVCDF function in PROC FCMP to generate a random sample from any continuous distribution.
- Compute limited moments of any distribution by using the LIMMOMENT function in PROC FCMP.
- Use a provided utility function to empirically separate the body and tail regions of a distribution.

Details

PROC SEVERITY fits the error model \( Y \sim F(\theta) \) for the response variable \( Y \), where \( F \) is a continuous probability distribution with parameters \( \theta \). You can use PROC SEVERITY to fit any distribution \( F \) by programatically defining its probability density function (PDF) and cumulative distribution function (CDF). You can also define several other functions to specify parameter constraints (bounds and fixed values), initial parameter values, and so on. Ten predefined distributions are provided.

The regression effects are modeled by extending the error model as

\[
Y \sim F(\theta, \Omega), \quad \theta = \theta_0 \exp \left( \sum_{j=1}^{k} \beta_j x_j \right)
\]

where \( \theta \) is the scale parameter of \( F \), \( \Omega \) is the set of nonscale parameters of \( F \), \( \theta_0 \) is the base value of \( \theta \), and \( \beta_j \) are the regression parameters that correspond to each of the \( k \) regressors, \( x_j \). To model regression effects, a distribution must have a scale parameter or a parameter \( p = \log(\theta) \). You can use all the predefined distributions, except Tweedie, to model regression effects.

The distribution and regression parameters of a model are estimated by minimizing either the negative of the log of the likelihood function or an objective function that you specify programatically.

You can specify losses that are unobserved due to truncation and inexact losses (censoring). The default estimation method accounts for such losses by maximizing the likelihood function,

\[
L = \prod_{i \in U} F_\theta(t_i^f) - F_\theta(t_i^l) \prod_{j \in C} F_\theta(c_j^l) - F_\theta(c_j^f)
\]

where \( U \) and \( C \) denote sets of uncensored and censored observations, respectively, \( y \) denotes the uncensored loss, \( (t_i^l, t_i^f) \) is the truncation interval, and \( (c_j^l, c_j^f) \) is the censoring interval. The estimation of the EDF also accounts for censored and truncated losses. If you specify both left- and right-censoring,
then Turnbull’s expectation maximization algorithm is used; otherwise, Kaplan-Meier’s (KM) product limit estimator, extended to account for truncation, is used. EDF standard errors are computed by using Greenwood’s formula (KM estimator) or the Hessian of Turnbull’s likelihood.

PROC SEVERITY computes likelihood-based fit statistics (AIC, AICC, BIC) and EDF-based fit statistics (Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling) for all models.

**Example: Analyzing Insurance Claims**

Let LClaims be a SAS data set that records the losses incurred by an auto insurance company. For each loss, the deductible and policy limit are also recorded. As an actuary of the company, you want to analyze whether the data follow a loglogistic or lognormal distribution. The following statements define the loglogistic distribution (llogis) and fit it along with the predefined lognormal distribution (logn):

```sas
proc fcmp library=sashelp.svrtdist outlib=t.mydist.severity;
function llogis_logpdf(x,Theta,Gamma);
  z = (x/Theta);
  return (log(Gamma) - log(Theta) + (Gamma-1)*log(z) - 2*log1px(z**Gamma));
endsub;
function llogis_logcdf(x,Theta,Gamma);
  z = (x/Theta);
  return (Gamma*log(z) - log1px(z**Gamma));
endsub;
subroutine llogis_parminit(n,x[*],nx[*],F[*],Ft,Theta,Gamma);
  outargs Theta, Gamma;
  q1=sqrtutil_percentile(0.25,n,x,F,Ft);
  q2=sqrtutil_percentile(0.75,n,x,F,Ft);
  Theta = exp((log(q1)+log(q2))/2);
  Gamma = 2*log(3)/(log(q2) - log(q1));
endsub;
quit;
```

By examining several outputs that are prepared by PROC SEVERITY (Figures 1–4), you can conclude that the loglogistic distribution fits the data well and is a better fit than the lognormal distribution.

**System Requirements**

Estimation using the SEVERITY procedure requires:
- BASE SAS® 9.3 or later
- SAS/ETS 9.3 or later

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2012, SAS Institute Inc. All rights reserved. 000000_0210