

Application of heavy-tailed distribution using PROC IML, NLMIXED, and SEVERITY

Palash Sharma, University of Kansas Medical Center, Kansas City, KS

John Keighley, Ph.D. University of Kansas Medical Center, Kansas City, KS

ABSTRACT

Probabilistic heavy-tailed distribution (Pareto, Weibull, Burr etc.) theory has vast applications involving in many real-life situations and natural phenomena. This area of research attracts not only for theoretical probabilistic nature but also for various branches of statistics. Heavy-tailed distribution are also used for modeling various biological, actuarial, financial, economic, hydrological, and engineering data. In this paper, we are modeling the cancer patient dataset and electrical blackouts dataset in the USA with generalized Pareto distribution. We also simulate the data arising from generalized Pareto distribution and estimate the parameters of generalized Pareto distribution (GPD) using maximum likelihood estimation. A suite of SAS procedure is used for all computation, specifically Procedure IML, SEVERITY, NLMIXED.

Keywords: MLE, Pareto distribution, GPD, heavy tail, NLMIXED etc.

INTRODUCTION

In recent years, heavy-tailed distribution has gained a lot of interest in literature. Heavy tail distribution receives much popularity for its flexible modeling capability and skewed shaped form. Therefore, it is extensively used for many applied sciences including economics, statistics, risk management and actuarial science and survival analysis, and lifetime data analysis. Generalization of Pareto distribution also exists in the various form in the literature. Generalized Pareto distribution (GPD) was first introduced by Pickands (1975) for modeling extreme value data. The Pareto distribution is a special case of generalized Pareto distribution. Pareto distribution is generated by compounding a heavy-tailed conditional gamma distribution with parameters k and θ^{-1} , where the weight for θ has a gamma distribution with parameters α and β . The extension of pareto distribution named Kumaraswamy Pareto distribution was first studied by Pereira et al (2012). Further, Akinsete et al. (2008) and Mahmoudi (2011) extended the Pareto and generalized Pareto distribution by introducing the beta Pareto and beta generalized Pareto distribution, respectively, which is based on a T-G family of distributions introduced by Eugene et al. (2002).

In this paper, we will present a statistical framework for discerning generalized Pareto distribution using empirical data. We will introduce Pareto distribution and generalized version of pareto distribution and demonstrate its flexibility by showing the wide variety of shapes of the distribution, cumulative density function, and reliability function. We will also estimate the unknown parameters of the generalized Pareto distribution using Maximum likelihood method using PROC IML and PROC NLMIXED. Two real data sets are considered for fitting generalized Pareto distribution along with the goodness of fit test using PROC SEVERITY in SAS procedure.

HEAVY TAIL PARETO DISTRIBUTION

Heavy-tailed distribution is probability distribution whose tails is bounded by exponentially. Mathematically, $\lim_{x \rightarrow \infty} e^{\lambda x} P(X > x) = +\infty$. They usually have heavier tails than the exponential family. Moreover, moment generating function for heavy tail distribution is not finite on the positive real line. The distribution $F(\cdot)$ from the random variable X is heavy tail if and only if

$$\int_{-\infty}^{\infty} e^{\lambda x} F(x) dx = \infty; \quad \forall, \lambda > 0.$$

The probability density function (PDF) for a random variable X is said to be a Pareto type II distribution or simply Pareto distribution, if it is given as $f_X(x) = \frac{\alpha \beta^\alpha}{(x+\beta)^{\alpha+1}}; \quad \forall, x > \beta$ and $\alpha, \beta > 0$.

The Pareto distribution is skewed to the right and characterized by a shape parameter α and a scale

parameter β . The density $f_X(x)$ is a decreasing function of x and achieves its maximum when x is smallest. The cumulative density function (CDF) of Pareto distribution can be written as $F_X(x) = 1 - \left(\frac{\beta}{x+\beta}\right)^\alpha$.

The hazard rate function/survival function can be written as in the form $h(x) = \frac{f_X(x)}{1-F_X(x)}$.

The PROC IML code for generating Pareto density plot is given in the appendix.

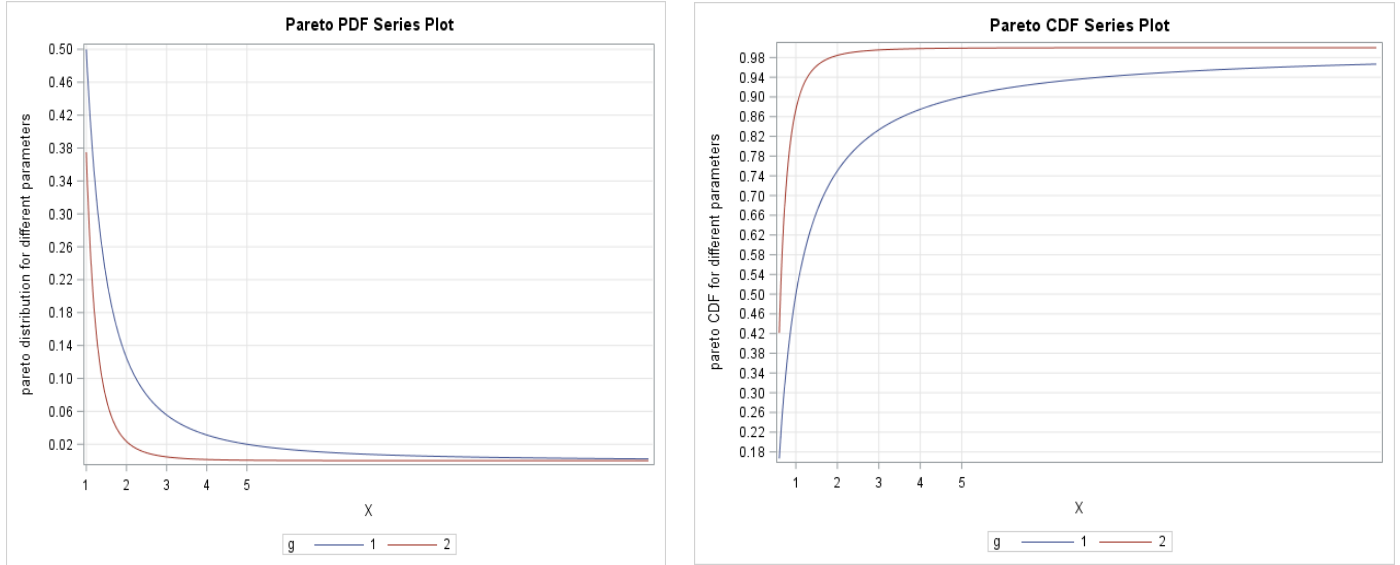


Figure 1: PDF and CDF of Pareto distribution for $\alpha = 1$ and $\beta = 0.5$ and $\alpha = 3$ and $\beta = 0.3$.

GENERALIZED PARETO DISTRIBUTION (GPD)

The probability density function of the three parameter Generalized Pareto distribution is

$$f_X(x) = \frac{1}{\sigma} \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-(1+\frac{1}{\xi})}, \quad x \geq \mu, \sigma \geq 0, \xi > 0 \quad (1)$$

Here μ is the location parameter, σ is the scale parameter and ξ is the shape parameter. In this paper, we are working with two parameter GPD model with $\mu = 0$. Therefore, two parameter GPD model is,

$$f_X(x) = \frac{1}{\sigma} \left(1 + \frac{\xi x}{\sigma}\right)^{-(1+\frac{1}{\xi})}, \quad x \geq 0, \sigma \geq 0, \xi > 0$$

The cumulative distribution function of the GPD can be written as,

$$F_X(x) = 1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}, \quad \sigma \geq 0, \xi > 0$$

The GPD model in equation (1) is generalized in the sense that it contains various special cases. When $\xi > 0$ and $\mu = 0$, the distribution is in the form of Pareto distribution with the transformation $\alpha = \frac{1}{\xi}$, $\beta = \frac{\sigma}{\xi}$. When $\xi = 0$, the GPD model is reduced to exponential distribution and When $\xi = 1$, the GPD model becomes uniform $U(0, \sigma)$.

The SAS code for generating PDF and CDF of generalized Pareto distribution is given in the appendix.

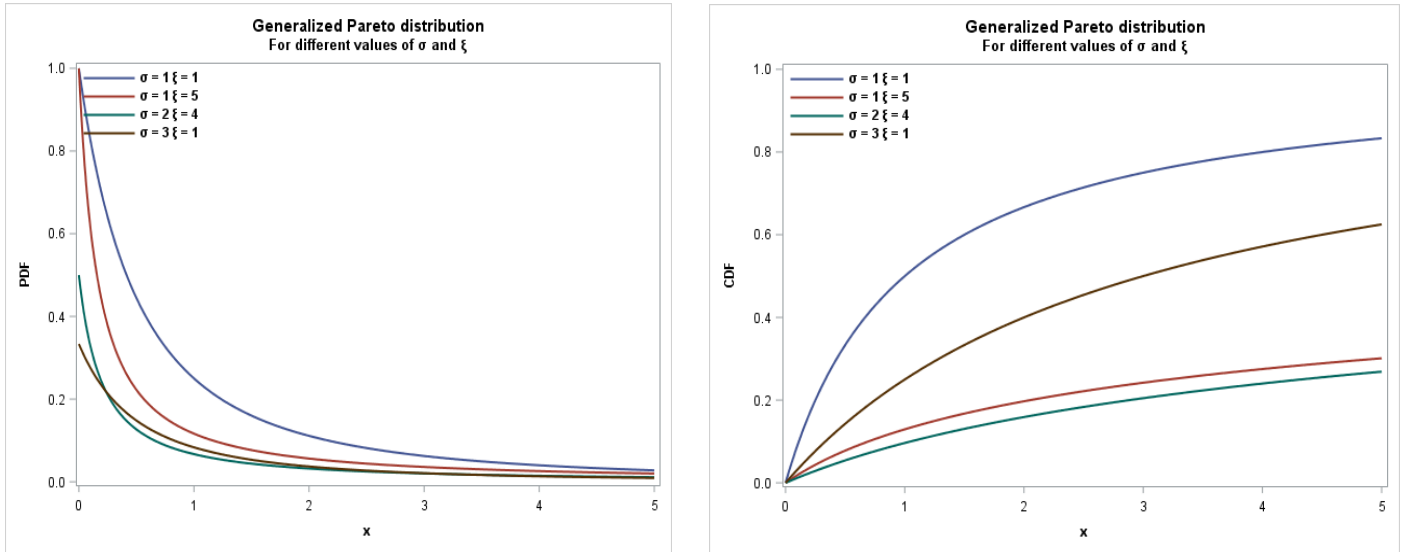


Figure 2: PDF and CDF of Generalized Pareto distribution for various values of σ and ξ .

SIMULATION TECHNIQUE

To simulate the data from generalized Pareto distribution, we use the probability integral transform theorem. By using the CDF of GPD and let it equal to U where, $U \sim \text{Uni}(0,1)$, we can get the following closed form expression for generating random variable arising from GPD model.

$$X = F^{-1}(U) = \frac{\sigma}{\xi} \left[\frac{1}{(1-U)^\xi} - 1 \right]$$

Here, U is the standard uniform random variable. The SAS code is given in the appendix to simulate data arising from Generalized Pareto distribution.

NUMERICAL OPTIMIZATION USING PROC NLMIXED

The estimation of α and β can be obtained using maximum likelihood estimation. The likelihood function of generalized Pareto distribution can be expressed as:

$$L(\alpha, \beta; x) = \prod_{i=1}^n \frac{1}{\sigma} \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-(1+\frac{1}{\xi})} = \frac{1}{\sigma^n} \prod_{i=1}^n \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-(1+\frac{1}{\xi})}$$

We consider Monte Carlo simulation technique to assess the performance of the MLEs of σ and ξ . We carry out the simulation using SAS software for various observations from generalized Pareto distribution. We consider different sample size $n=50, 500, 800, 1000$ and the number of replications is 5000. For simulation purpose, we are considering the true parameters value for the generalized Pareto distribution is $\sigma = 2$ and $\xi = 3$. Table 1 below shows the estimated parameters of σ and ξ , standard error, and confidence interval for the parameters. We observe that, when the sample size is increasing, the standard error for parameters α and β are very small or close to zero and converges to true parameter value.

Table1: Standard error and estimated parameter when is $\sigma = 2$ and $\xi = 3$

| n | Parameters | Value | S.E. | 95% Confidence Limits |
|------|----------------|--------|----------|-----------------------|
| 50 | $\hat{\sigma}$ | 2.0062 | 0.01133 | 1.9840 2.0284 |
| | $\hat{\xi}$ | 2.9898 | 0.007977 | 2.9742 3.0054 |
| 500 | $\hat{\sigma}$ | 2.0052 | 0.003587 | 1.9982 2.0123 |
| | $\hat{\xi}$ | 2.9984 | 0.002529 | 2.9935 3.0034 |
| 800 | $\hat{\sigma}$ | 2.0028 | 0.002833 | 1.9972 2.0083 |
| | $\hat{\xi}$ | 2.9997 | 0.002000 | 2.9958 3.0036 |
| 1000 | $\hat{\sigma}$ | 2.0010 | 0.002532 | 1.9960 2.0059 |
| | $\hat{\xi}$ | 3.0004 | 0.001789 | 2.9969 3.0039 |

APPLICATION OF GPD MODEL USING PROC SEVERITY

Example 1:

The dataset consists of cancer patient data. The data represent the remission times (in months) of a random sample of 128 bladder cancer patients Lee and Wang (2003). Estimation of the parameters of GP distribution is done by PROC SEVERITY and compare it with other predefined distributions. We use Akaike information criterion (AIC), Consistent Akaike information criterion (AICC), Bayesian information criterion (BIC) is used to select the best model to fit the cancer data. Anderson darling and KS statistics is also given for goodness of fit test. Plot of the fitted densities and histogram, observed versus predicted probabilities for cancer patient data also given in figure 3 and 4. We also compare the GPD model with other distributions such as inverse gaussian distribution, Log normal distribution, pareto distribution, Weibull distribution. Comparing with other univariate distributions, we can observe that generalized pareto distribution has the lowest AIC, AICC and BIC value. Also, P-P plot suggested that the data are approximately normally distributed. By inspection, GPD model is the best model to fit the cancer data.

Model Selection

| Distribution | Converged | -2 Log Likelihood | Selected |
|--------------|-----------|-------------------|----------|
| Gpd | Yes | 827.65117 | Yes |
| Igauss | Yes | 880.57157 | No |
| Logn | Yes | 830.19194 | No |
| Pareto | Yes | 827.65117 | No |
| Weibull | Yes | 828.16494 | No |

All Fit Statistics

| Distribution | -2 Log Likelihood | AIC | AICC | BIC | KS | AD |
|--------------|-------------------|-------------|-------------|-------------|-----------|-----------|
| Gpd | 827.65117 * | 831.65117 * | 831.74717 * | 837.35523 * | 1.11074 | 1.42650 |
| Igauss | 880.57157 | 884.57157 | 884.66757 | 890.27563 | 2.19822 | 7.86204 |
| Logn | 830.19194 | 834.19194 | 834.28794 | 839.89600 | 0.71672 * | 0.86238 * |
| Pareto | 827.65117 | 831.65117 | 831.74717 | 837.35523 | 1.11076 | 1.42654 |
| Weibull | 828.16494 | 832.16494 | 832.26094 | 837.86900 | 0.80831 | 1.02110 |

Note: The asterisk (*) marks the best model according to each column's criterion.

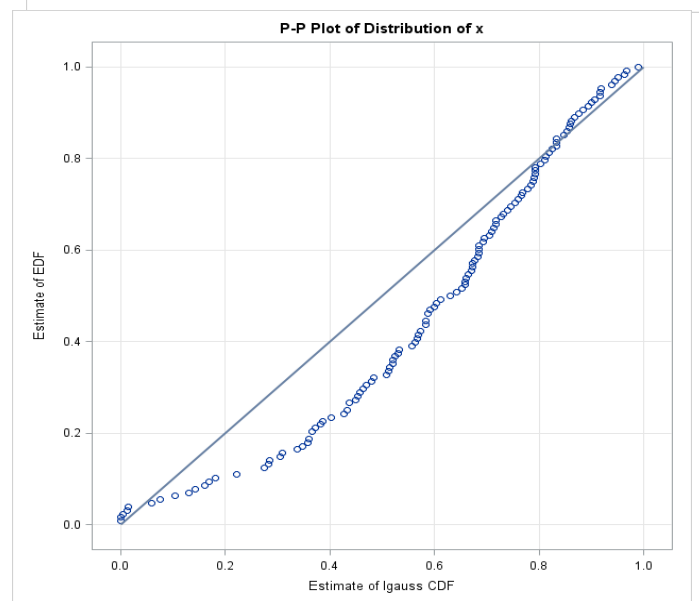
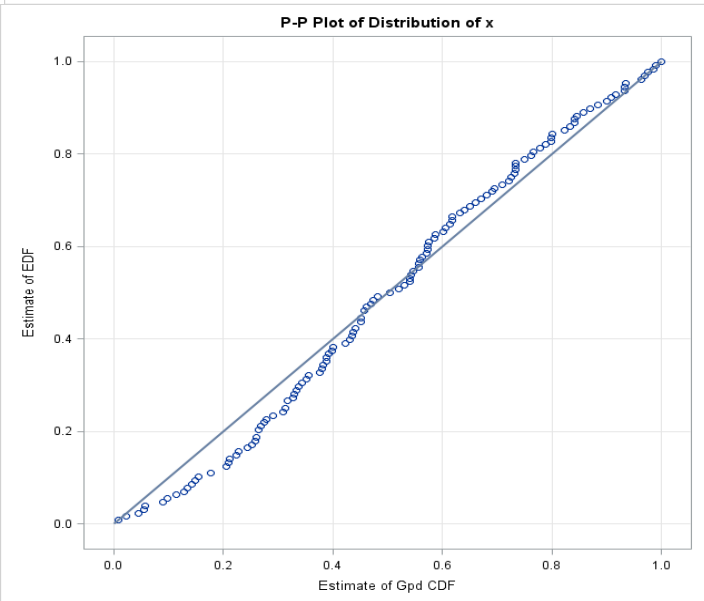
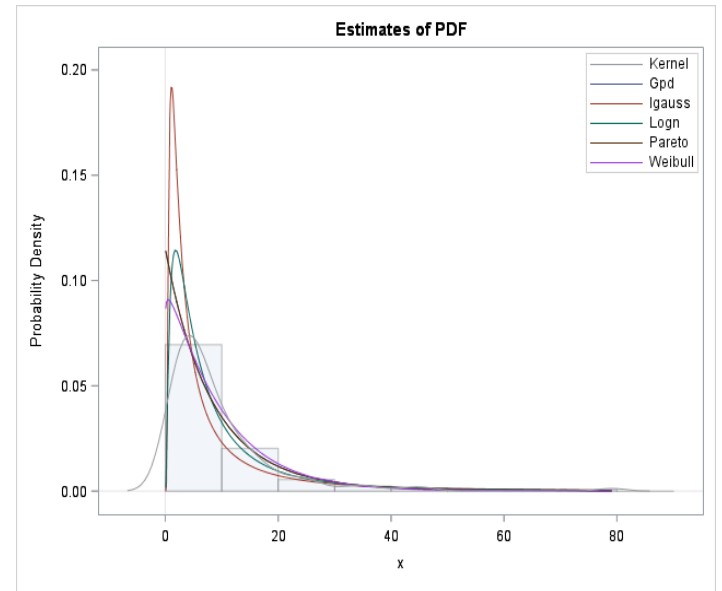
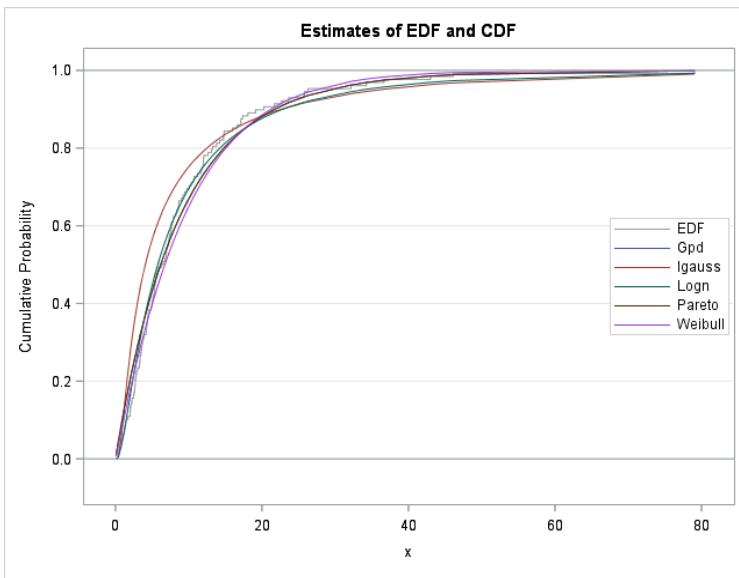


Figure 3: Cancer patient data fitting with various distribution.

Example 2:

In this dataset, we will investigate the number of customers affected by electrical blackouts in the United States between 1984 and 2002. The dataset was collected from Clauset et al. (2009). The disturbances dataset (NERC data) comes from the Disturbance Analysis Working Group (DAWG) Database who are investigating the disturbance occurred on the electric utilities in the USA. It is deemed to be the best-recorded blackout sources in the United States power transmission system. The measurement unit (blackout size), is measured by three different quantities, the amount of power loss, the number of customers affected and the restoration time (Minute) (10). We are modeling the number of customers affected by blackouts and fit the dataset by various distributions using PROC SEVERITY.

| All Fit Statistics | | | | | | | | | | | | |
|--------------------|-------------------|--------|--------|--------|-----------|-----------|-----------|--|--|--|--|--|
| Distribution | -2 Log Likelihood | AIC | AICC | BIC | KS | AD | CvM | | | | | |
| Gpd | 5549 * | 5553 * | 5553 * | 5560 * | 0.73019 * | 0.96456 * | 0.08680 * | | | | | |
| Burr | 5554 | 5560 | 5560 | 5570 | 0.92229 | 1.47112 | 0.19399 | | | | | |
| Pareto | 5550 | 5554 | 5554 | 5561 | 0.86627 | 1.10749 | 0.13371 | | | | | |
| Weibull | 5579 | 5583 | 5583 | 5590 | 1.16498 | 3.46025 | 0.48864 | | | | | |
| Exp | 5644 | 5646 | 5646 | 5650 | 3.13880 | 19.50451 | 3.79835 | | | | | |
| Gamma | 5605 | 5609 | 5609 | 5615 | 1.82088 | 6.39766 | 1.15394 | | | | | |

Note: The asterisk (*) marks the best model according to each column's criterion.

| Estimation Details | | | | | | |
|--------------------|-----------|------------|----------------|------------------|-----------------|----------------|
| Distribution | Converged | Iterations | Function Calls | Gradient Updates | Hessian Updates | Time (Seconds) |
| Gpd | Yes | 29 | 68 | 527 | 495 | 0.00 |
| Burr | Maybe | 100 | 213 | 5355 | 5252 | 0.02 |
| Pareto | Maybe | 100 | 264 | 5355 | 5252 | 0.00 |
| Weibull | Yes | 36 | 85 | 779 | 740 | 0.00 |
| Exp | Yes | 0 | 4 | 5 | 2 | 0.00 |
| Gamma | Yes | 3 | 16 | 20 | 14 | 0.00 |

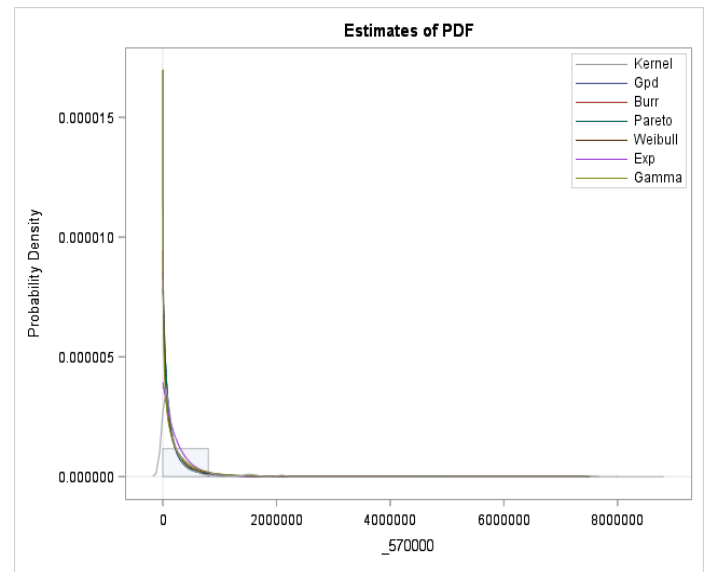
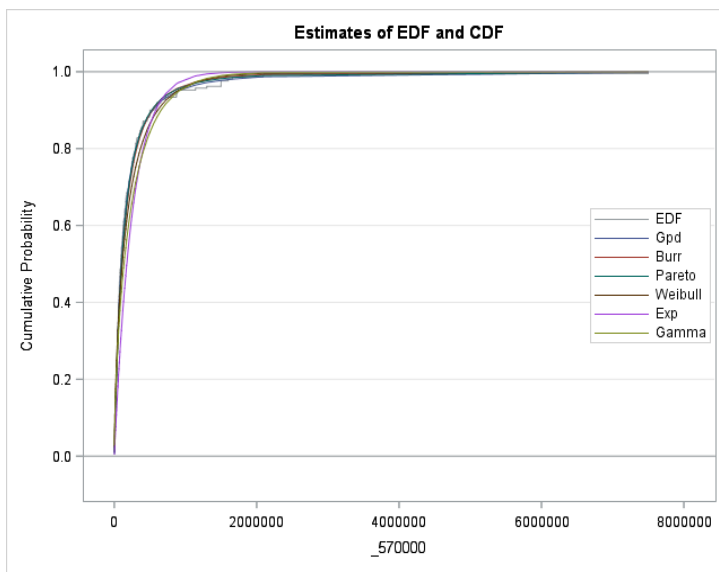


Figure 4: customer blackout data modeling by various distributions.

We compare our dataset with Burr distribution, Pareto distribution, Weibull distribution, Gamma and Exponential distributions. We consider criteria like log-likelihood (LL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and AICC to identify the better model to fit the data. The better distribution is corresponding to the smaller values of LL, AIC, AICC, and BIC. Also, the Kolmogorov–Smirnov (K–S) goodness-of-fit statistic and Anderson darling test statistics are used to conduct the appropriateness of the model. We observe that generalized Pareto distribution gives the smallest value regardless of AIC, AICC or BIC. It gives the better fit to our dataset. Also, from the P-P plot and cumulative density plot vs the observed data gives the similar conclusion. Looking at the information criteria, we can see that, other distributions do not provide an adequate fit to the dataset.

CONCLUSION

We define and study the Pareto and generalized version of Pareto distribution. Various properties of the distribution are also investigated. The method of maximum likelihood is suggested using SAS NLMIXED. In both applications, generalized Pareto distribution provides significantly better fit than the other distributions. Further extension of the generalized Pareto model and the application in different scenarios will be investigated in future.

REFERENCES

- SAS® Institute Inc 2016. SAS® Enterprise Miner and SAS® text Miner Procedures Reference for SAS 9.4®. Cary, NC: SAS Institute Inc.
- Aban, I.B., Meerschaert, M.M. and Panorska, A.K. 2006. "Parameter estimation for the truncated Pareto distribution", *J. Am. Statist. Assoc.* 101(473) pp. 270–277.
- Akinsete, A., Famoye, F. and Lee, C. 2008. "The beta-Pareto distribution." *Statistics*, 42(6), 547-563.
- Chen, J., Thorp, J. S. and Parashar, M. 2001, Analysis of electric power disturbance data. *In 34th Hawaii International Conference on System Sciences, IEEE Computer Society.*
- Clauset, A, Shalizi, C.R. and Newman, M.E.J.2009. "Power-law distributions in empirical data" *SIAM Review* 51(4), 661-703. ([arXiv:0706.1062](https://arxiv.org/abs/0706.1062)).
- Eugene, N., Lee, C. and Famoye, F. 2002. "The beta-normal distribution and its applications," *Commun. Statist. Theory Meth.*, pp. 497–512.
- Levy, M. and Levy, H. 2003. "Investment talent and the Pareto wealth distribution: Theoretical and experimental analysis," *Rev. Econ. Statist.* 85(3), pp. 709–725.
- Lee, E. T. and Wang, J. 2003. "Statistical methods for survival data analysis". *volume 476. John Wiley & Sons.*
- Mahmoudi, E. 2011. "The beta generalized Pareto distribution with application to lifetime data." *Mathematics and Computers in Simulation*, 81, 11, 2414–2430.
- Newman, M.E.J. 2005. Power laws, Pareto distributions and Zipf's law, *Contemp. Phys.* 46(5), pp. 323–351.
- Pereira, M. B., Silva, R. B., Zea, L. M., and Cordeiro, G. M. 2012. "The Kumaraswamy Pareto distribution". *arXiv preprint arXiv:1204.1389*.
- Pickands, j. 1975. "Statistical inference using extreme order statistics", *Ann. Statist.* 3(1), pp. 119–131.

ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. Jonathan D. Mahnken and Dr. John Keighley for giving me some insight of the SAS code.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Palash Sharma

Department of Biostatistics

University of Kansas Medical Center

3905 Rainbow Boulevard, Kansas City, KS

Email: psharma4@kumc.edu

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.

APPENDIX

```
/*## Pareto pdf for different parameters.*/

proc iml;
start paretoPDF(x, a, b);
f=j(nrow(x),ncol(x),0);
idx = loc(x>b);
if ncol(idx)>0 then do;
v=x[idx];
f[idx] = (a*b##a)#(1/(x##(a+1)));
end;
return(f);
finish;
x=do(1,15,0.01);
y1 = paretoPDF(x,1,0.5);
y2 = paretoPDF(x,3,0.5);
g = repeat({1,2}, 1, ncol(x)); /* 1,1,1,...,2,2,2 */
x = x || x ;
y = y1 || y2;
title "Pareto distribution Series Plot ";
call Series(x, y) group=g
            grid={X Y}
            label={"X" "pareto distribution for different parameters"}
            xvalues=0:5
            yvalues=do(0,0.5,0.02);

/*Plotting density function of generalized pareto distribution*/
%let sigma_1=1;
%let zeta_1 = 1;

%let sigma_2=1;
%let zeta_2 = 5;

%let sigma_3=2;
%let zeta_3 = 4;

%let sigma_4=3;
%let zeta_4 = 1;

data GPD;
do x = 0 to 5 by 0.01;
pdf_gpd_1= (1/&sigma_1.)*(1+&zeta_1. * x )**(-1-(1/&zeta_1.));
pdf_gpd_2= (1/&sigma_2.)*(1+&zeta_2. * x )**(-1-(1/&zeta_2.));
pdf_gpd_3= (1/&sigma_3.)*(1+&zeta_3. * x )**(-1-(1/&zeta_3.));
pdf_gpd_4= (1/&sigma_4.)*(1+&zeta_4. * x )**(-1-(1/&zeta_4.));
output;
end;
run;

title "Generalized Pareto distribution";
title2 "For different values of (*ESC*){unicode sigma} and (*ESC*){unicode xi} ";

proc sgplot data = GPD noautolegend;
series x=x y=pdf_gpd_1 / lineattrs = (thickness=2)
legendlabel=(*ESC*){unicode sigma} = 1 (*ESC*){unicode xi} = 1";
```

```

series x=x y=pdf_gpd_2 / lineattrs = (thickness=2)
legendlabel="(*ESC*){unicode sigma} = 1 (*ESC*){unicode xi} = 5";
series x=x y=pdf_gpd_3 / lineattrs = (thickness=2)
legendlabel="(*ESC*){unicode sigma} = 2 (*ESC*){unicode xi} = 4";
series x=x y=pdf_gpd_4 / lineattrs = (thickness=2)
legendlabel="(*ESC*){unicode sigma} = 3 (*ESC*){unicode xi} = 1";
keylegend / position=NW location=inside across=1 noborder valueattrs=(size=9
weight=bold);
yaxis min=0 max=1 label="PDF" labelattrs=(size=9 weight=bold);
xaxis min=0 max=5 label="x" labelattrs=(size=9 weight=bold);
run;
title;
/*## Plotting CDF for PARETO DISTRIBUTION*/

proc iml;
start paretoCDF(x,a,b);
f=j(nrow(x),ncol(x),0);
idx=loc(x>b);
if ncol(idx)>0 then do;
v=x[idx];
f[idx] = 1-(b/x)##a;
end;
return(f);
finish;

x=do(0.6,15,0.01);
y1 = paretoCDF(x,1,0.5);
y2 = paretoCDF(x,3,0.5);
g = repeat({1,2}, 1, ncol(x)); /* 1,1,1,...,2,2,2 */
x = x || x;
y = y1 || y2;
title "Pareto CDF Series Plot ";
call Series(x, y) group=g
grid={X Y}
label={"X" "pareto CDF for different parameters"}
xvalues=0:5
yvalues=do(0,0.5,0.02);

/*Plotting cumulative density function of generalized pareto distribution*/

%let zeta_1 = 1;
%let zeta_2 = 5;
%let zeta_3 = 9;
%let zeta_4 = 3;

data GPD_cdf;
do x = 0 to 5 by 0.001;
cdf_gpd_1= 1-(1+&zeta_1.*x)**((-1/&zeta_1.));
cdf_gpd_2= 1-(1+&zeta_2.*x)**((-1/&zeta_2.));;
cdf_gpd_3= 1-(1+&zeta_3.*x)**((-1/&zeta_3.));;
cdf_gpd_4= 1-(1+&zeta_4.*x)**((-1/&zeta_4.));;
output;
end;
run;

```

```

title "Generalized Pareto distribution";
title2 "For different values of (*ESC*){unicode theta} and (*ESC*){unicode xi}
";

proc sgplot data = GPD_cdf noautolegend;
series x=x y=cdf_gpd_1 / lineattrs = (thickness=2) legendlabel="(*ESC*){unicode
theta} = 1 (*ESC*){unicode xi} = 1";
series x=x y=cdf_gpd_2 / lineattrs = (thickness=2) legendlabel="(*ESC*){unicode
theta} = 1 (*ESC*){unicode xi} = 5";
series x=x y=cdf_gpd_3 / lineattrs = (thickness=2) legendlabel="(*ESC*){unicode
theta} = 2 (*ESC*){unicode xi} = 20";
series x=x y=cdf_gpd_4 / lineattrs = (thickness=2) legendlabel="(*ESC*){unicode
theta} = 3 (*ESC*){unicode xi} = 1";
keylegend / position=NW location=inside across=1 noborder valueattrs=(size=9
weight=bold);
yaxis min=0 max=1 label="PDF" labelattrs=(size=9 weight=bold);
xaxis min=0 max=1 label="x" labelattrs=(size=9 weight=bold);
run;
title;
/*## plotting hazard rate function of Pareto distribution*/

proc iml;
start paretosurv(x,a,b);
f=j(nrow(x),ncol(x),0);
idx=loc(x>b);
if ncol(idx)>0 then do;
v=x[idx];
f[idx] = ((a*b##a)#(1/(x##(a+1))))/((b/x)##a);
end;
return(f);
finish;

x=do(0.6,15,0.01);
y1 = paretosurv(x,1,0.5);
y2 = paretosurv(x,3,0.5);
g = repeat({1,2}, 1, ncol(x)); /* 1,1,1,...,2,2,2 */
x = x || x;
y = y1 || y2;
title "Pareto Survival Series Plot ";
call Series(x, y) group=g
grid={X Y}
label={"X" "pareto survival plot for different parameters"}
xvalues=0:5
yvalues=do(0,0.5,0.02);

title;

proc univariate data=pareto;
var x;
histogram;
inset n mean(5.3) std='std Dev'(5.3) skewness(5.3);
output out=moments mean=_mean_ std=_stdev_
skewness=_skewness_ kurtosis=_kurtosis_ n=_count_;
run;

```

```

data moments;
set moments;
a_mme = _mean_*( _stdev_**2+_mean_**2)/
        (_stdev_**2-_mean_**2);
b_mme = 2*( _stdev_**2)/(_stdev_**2 - _mean_**2);
run;

data _null_;
set moments;
call symput("a_mme",a_mme);
call symput("b_mme",b_mme);
run;
/**/
/*## MLE estimation */
proc iml;
use pareto;
read all into x_obs;
n_obs = nrow(x_obs);

b_o = &b_mme;
a_o = &a_mme;

val_o= b_o//a_o;

do i=1 to 10000;

/*First derivative*/
d_b = n_obs/b_o + n_obs*log(a_o) - sum(log(x_obs + j(n_obs,1,a_o)));
d_a = n_obs*b_o /a_o - (b_o+1) * sum((x_obs + j(n_obs,1,a_o))##(-1));

/*Second derivative*/

d1_b = -n_obs/(b_o**2);
d1_ab = n_obs/a_o - sum(x_obs+j(n_obs,1,a_o));
d1_a = -n_obs*b_o/(a_o**2) + (b_o+1) * sum((x_obs + j(n_obs,1,a_o))##(-2));

par_1 = d_b // d_a;
par_2 = (d1_b || d1_ab) // (d1_ab || d1_a);

val_n = val_o - (inv(par_2))*par_1;
diff = sum(abs(val_n-val_o));
b_o = val_n[1];
a_o = val_o[2];

val_o = val_n;

if diff < 0.0001 then i =10000;
end;
print diff val_n;
quit;

```

```

/*Simulation Technique to generate GPD random variable*/

%let sigma=2;
%let zeta=3;
%let n=800;
%let nsample=5000;

data simulation(drop=i u);
do sampleid = 1 to &nsample.;
do i= 1 to &n.;
u=ranuni(1234);
x=-(&sigma./&zeta.)*(1-(1/(1-u)**(&zeta.)));
output;
end;
end;
run;

/*Maximum likelihood estimation using proc NLMIXED*/

proc nlmixed tech=trureg data=simulation MAXFUNC=900 MAXITER=900;
parms zeta=2.5 sigma=1.5;
title 'Generalized pareto distribution ';
bounds 0 < sigma zeta;
xx=(1/sigma)*(1+(zeta*x)/sigma)**(-(1+(1/zeta)));
ll=log(xx);
model x~general(ll);
run;

/* Data Analysis;*/

data cancer;
input x @@;
cards;
0.08 2.09 3.48 4.87 6.94 8.66 13.11 23.63 0.20 2.23
3.52 4.98 6.97 9.02 13.29 0.40 2.26 3.57 5.06 7.09
9.22 13.80 25.74 0.50 2.46 3.64 5.09 7.26 9.47 14.24
25.82 0.51 2.54 3.70 5.17 7.28 9.74 14.76 26.31 0.81
2.62 3.82 5.32 7.32 10.06 14.77 32.15 2.64 3.88 5.32
7.39 10.34 14.83 34.26 0.90 2.69 4.18 5.34 7.59 10.66
15.96 36.66 1.05 2.69 4.23 5.41 7.62 10.75 16.62 43.01
1.19 2.75 4.26 5.41 7.63 17.12 46.12 1.26 2.83 4.33
5.49 7.66 11.25 17.14 79.05 1.35 2.87 5.62 7.87 11.64
17.36 1.40 3.02 4.34 5.71 7.93 11.79 18.10 1.46 4.40
5.85 8.26 11.98 19.13 1.7 3.25 4.50 6.25 8.37 12.02
2.02 3.31 4.51 6.54 8.53 12.03 20.28 2.02 3.36 6.76
12.07 21.73 2.07 3.36 6.93 8.65 12.63 22.69
;
run;
/**/
/*http://support.sas.com/resources/papers/proceedings15/3307-2015.pdf*/

* Close all open ODS destinations;
ods _all_ close;
options nocenter nodate nonumber;
title1 'Table [your table number here]';
title2 '[your table title]';

```

```
ods rtf file='APA.rtf' bodytitle_aux style=Journal;

proc severity data=cancer print=all plots=pp;
  loss x;
  dist gpd IGAUSS logn pareto weibull;
  nloptions absfconv=2.0e-5 maxiter=100 maxfunc=500;
run;
quit;
ods rtf close;

/*Data Analysis 2 */

proc severity data=work.pal print=all plots=pp;
  loss x;
  dist burr pareto weibull exp gamma;
  nloptions absfconv=2.0e-5 maxiter=100 maxfunc=500;
run;
```