

Appendix E

Constructing a Moment-Ratio Diagram in SAS

Contents

E.1	Creating Annotation Data Sets for the SGPLOT Procedure	367
E.2	Creating a Moment-Ratio Diagram	367
E.3	Displaying a Moment-Ratio Diagram	371
E.4	References	372

E.1 Creating Annotation Data Sets for the SGPLOT Procedure

This supplemental material describes how to use the annotation facility for the SGPLOT procedure to create a moment-ratio diagram. Chapter 16, “Moment Matching and the Moment-Ratio Diagram,” of *Simulating Data with SAS* (Wicklin 2013) describes how to use a moment-ratio diagram to understand the relationships between distributions and to display the variation in the skewness and kurtosis statistics for simulated samples.

Beginning with SAS 9.3, the SGPLOT procedure supports the display of annotations. Annotations are created by using a special data set, called the annotation data set. To create a moment-ratio diagram, perform the following steps:

1. Create an annotation data set that draws the moment-ratio diagram. The moment-ratio diagram in this supplemental material is for continuous distributions. See Vargo, Pasupathy, and Leemis (2010) and Chapter 16 of Wicklin (2013) for a description of a moment-ratio diagram.
2. Create a SAS data set that contains the skewness and kurtosis values to be plotted.
3. Use the SGANNO= option on the PROC SGPLOT statement to display the moment-ratio diagram overlaid with a scatter plot of the skewness and kurtosis values.

E.2 Creating a Moment-Ratio Diagram

See Heath (2011) for an overview of using annotation data sets with the statistical graphics procedures. The annotate data set in this section creates a moment-ratio diagram with the following features:

368 Appendix E: Constructing a Moment-Ratio Diagram in SAS

- a region for the beta distribution
- curves for the lognormal and gamma distributions
- points for the normal, the exponential, and the t distributions with 5–12 degrees of freedom

In the following DATA steps, each region or curve is created in its own annotation data set. Then the various annotation data sets are concatenated to create the Anno data set, which contains all of the annotations. The final moment-ratio diagram is shown in Figure E.1.

The following SAS statements define macros that specify the range of skewness and kurtosis values for the moment-ratio diagram:

```
/* Create the annotation data set for the moment-ratio diagram */
/* All computations are in terms of EXCESS kurtosis.
   Use Y2AXIS to add FULL kurtosis axis. */
%let xL = -2.4; %let xR = 2.4;      /* range of skewness is [xL, xR] */
%let yB = -2.0; %let yT = 10.0;   /* range of kurtosis is [yB, yT] */
%let yB2 = %sysevalf(&yB+3);      /* range for full kurtosis      */
%let yT2 = %sysevalf(&yT+3);

/* Macro to convert between full kurtosis and excess kurtosis */
%macro Ex2Full(k); ((&k)+3) %mend;
%macro Full2Ex(k); ((&k)-3) %mend;
```

The following DATA step creates an annotation data set for the boundary of the feasible region, $\kappa = 1 + \gamma^2$, where κ is the full kurtosis and γ is the skewness:

```
/* annotation data set for BOUNDARY of feasible (skew, kurt) region */
data Boundary(drop=x);
length function $12 Curve $12 LineColor $20;
retain DrawSpace "DataValue"
       LineColor "Black"
       Curve "Boundary";
function = "POLYLINE";
x1=&xL; y1=1+x1**2; y1=%Full2Ex(y1); output;
function = "POLYCONT";
do x = &xL to &xR by 0.1;
    x1=x; y1=1+x1**2; y1=%Full2Ex(y1);
    output;
end;
run;
```

The following DATA step draws the region of the moment-ratio diagram that corresponds to the (γ, κ) values for the beta family of distributions:

```

/* annotation data set for BETA region of moment-ratio diagram */
data Beta(drop=x);
length function $12 Curve $12 Label $24 LineColor $20;
retain DrawSpace "DataValue"
      Display "All"
      LineColor "LightGray"
      Curve "Beta";
function = "TEXT"; Anchor="Left ";
label = "Beta";
x1=&xL; y1=2+x1**2; y1=%Full2Ex(y1); output;
Transparency = 0.5; FillTransparency = 0.5;
FillColor = "LightGray";
function = "POLYGON";
x1=&xL; y1=1+x1**2; y1=%Full2Ex(y1); output;
function = "POLYCONT";
n=1;
do x = &xL to &xR+0.05 by 0.1;
  x1=x; y1=1+x1**2; y1=%Full2Ex(y1);
  output;
end;
do x = &xR to &xL-0.05 by -0.1;
  x1=x; y1=3+1.5*x1**2; y1=%Full2Ex(y1);
  output;
end;
run;

```

The following DATA step draws a curve on the moment-ratio diagram that corresponds to the (γ, κ) values for the lognormal family of distributions:

```

/* annotation data set for LOGNORMAL curve of moment-ratio diagram */
data LogNormal;
length function $12 Label $24 Curve $12 LineColor $20;
retain DrawSpace "DataValue"
      LineColor "DarkGreen"
      Curve "LogNormal";
function = "TEXT"; Anchor="Right";
label = "LogN";
drop var var0;
var0 = 0.355; var=var0;
x1= (exp(var)+2)*sqrt(exp(var)-1);
y1 = exp(4*var) + 2*exp(3*var) + 3*exp(2*var) - 6;
output;
function = "POLYLINE"; Anchor=" ";
output;
function = "POLYCONT";
do var = Var0 to 0.005 by -0.005;
  x1= (exp(var)+2)*sqrt(exp(var)-1);
  y1 = exp(4*var) + 2*exp(3*var) + 3*exp(2*var) - 6;
  output;
end;
x1=0; y1=0;output;
run;

```

370 Appendix E: Constructing a Moment-Ratio Diagram in SAS

The following DATA step draws a curve on the moment-ratio diagram that corresponds to the (γ, κ) values for the gamma family of distributions:

```

/* annotation data set for GAMMA curve of moment-ratio diagram */
data GammaCurve;
retain DrawSpace "DataValue";
length function $12 Label $24 Curve $12 LineColor $20;
retain LineColor "Magenta" Curve "Gamma";
drop a a0 dx N i;
function = "TEXT"; Anchor="Right";
label = "Gam";
a0 = (2/&xR)**2;
a = a0;
x1 = 2/sqrt(a); y1= 6/a; output;
Transparency = 0.5;
function = "POLYLINE"; Anchor=" "; output;
dx=0.1; N=floor(&xR/dx);
function = "POLYCONT";
do i = 2 to N;
    a = (2/(2/sqrt(a) - 0.1))**2;
    x1 = 2/sqrt(a); y1= 6/a;
    if (&xL<= x1 <=&xR) & (&yB <= y1 <= &yT) then output;
end;
x1=0; y1=0; output;
run;

```

The following DATA step draws points on the moment-ratio diagram that correspond to the (γ, κ) values for the exponential, normal, Gumbel, and several t distributions:

```

/* Annotation data set for special points in the moment-ratio diagram:
   Points for the exponential, normal, Gumbel, and t distribution with
   DOF 5,6,7,8,10,11,12. Also text that marks the invalid region. */
data MRPoints(drop=nu);
retain DrawSpace "DataValue";
length function $12 Label $24 Curve $12;
function = "TEXT";
Label="E"; x1 = 2; y1= 6; Curve="Exponential"; output;
Label="N"; x1 = 0; y1= 0; Curve="Normal"; output;
Label="G"; x1 = 1.14; y1= 2.4; Curve="Gumbel"; output;
Curve="T";
do nu=5 to 8;
    Label=cats("T",nu); x1=0; y1 = 6/(nu-4); output;
end;
do nu=10 to 12; /* plot ellipses marks */
    Label="."; x1=0; y1 = 6/(nu-4); output;
end;
Curve="Region";
do x1=-2,2;
    function="TEXT"; Label="Invalid"; y1=-1; output;
    function="TEXTCONT"; Label="Region"; output;
end;
run;

```

The following DATA step concatenates the previous annotation data sets into a single annotation data set. It is this data set that is used to draw the moment-ratio figures in Chapter 16 of Wicklin (2013).

```
/* Concatenate the pieces of the moment-ratio diagram */
data Anno;
  set Beta Boundary LogNormal GammaCurve MRPoints;
run;
```

E.3 Displaying a Moment-Ratio Diagram

After the annotation data set is created, you can overlay a scatter plot of skewness and kurtosis. The following SAS statements simulate 100 samples of size $N = 50$ from a Gamma(4) distribution. The MEANS procedure computes the skewness and kurtosis for each sample.

```
/* simulate 100 samples of size N=50 from a Gamma(4) distribution */
%let N = 50;
%let NumSamples = 100;
data Gamma(keep=x SampleID);
call streaminit(12345);
do SampleID = 1 to &NumSamples;
  do i = 1 to &N;
    x = rand("Gamma", 4);
    output;
  end;
end;
run;

/* compute the skewness and kurtosis of each sample */
proc means data=Gamma noprint;
  by SampleID;
  var x;
  output out=MomentsGamma skew=Skew kurt=Kurt;
run;
```

You can use the SGPLOT procedure to plot a scatter plot of the sample moments and use the SGANNO= option to display the moment-ratio diagram in the background. To make it easier to use, the following statements define a SAS macro that takes a data set, an annotation data set, and a transparency parameter for the scatter plot. This macro is not strictly necessary, and you can use the SGPLOT procedure directly for more control over the axes ranges.

```
/* Plot the moment-ratio diagram for a given data set and a
   given annotation data set. By default, the (skew, kurt) scatter plot
   is fully opaque, but you can set the transparency in the macro call. */
%macro PlotMRDiagram(DS, annoDS, Transparency=0);
  /* Given a data set that contains variables KURT and SKEW,
     add the FULLKURTOSIS variable and labels for the three variables. */
data &ds;
  set &ds;
  FullKurt = Kurt+3;
  label Kurt="Excess Kurtosis" FullKurt="Full Kurtosis" Skew="Skewness";
run;
```

```

proc sgplot data=&DS sganno=&annoDS noautolegend;
  scatter x=Skew y=Kurt / transparency=&Transparency;
  scatter x=Skew y=FullKurt / y2axis transparency=1; /* invisible */
  refline 0 / axis=x transparency=0.2;
  xaxis grid values=(-2.5 to 2.5 by 0.5);
  yaxis reverse grid values=(-2 to 10);
  y2axis reverse grid values=( 1 to 13);
run;
%mend PlotMRDiagram;

```

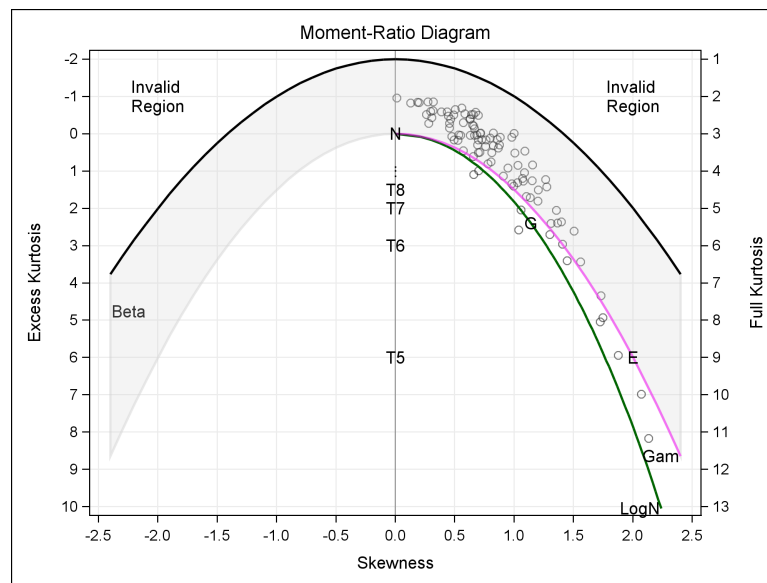
When you run the macro, a moment-ratio diagram is created. The data set must contain variables SKEW and KURT that contain the sample skewness and excess kurtosis, respectively. To see how the macro works, call it with the MomentsGamma data set and the Anno annotation data set:

```

/* Display a moment-ratio diagram in the background of a scatter
   plot of the sample skewness and kurtosis for simulated samples. */
title "Moment-Ratio Diagram";
%PlotMRDiagram(MomentsGamma, Anno, Transparency=0.5);

```

Figure E.1 Moment-Ratio Diagram for Simulated Gamma(4) Samples



E.4 References

- Heath, D. (2011), "Now You Can Annotate Your Statistical Graphics Procedure Graphs," in *Proceedings of the SAS Global Forum 2011 Conference*, Cary, NC: SAS Institute Inc.
 URL <http://support.sas.com/resources/papers/proceedings11/277-2011.pdf>
- Vargo, E., Pasupathy, R., and Leemis, L. M. (2010), "Moment-Ratio Diagrams for Univariate Distributions," *Journal of Quality Technology*, 42, 1–11.
- Wicklin, R. (2013), *Simulating Data with SAS*, Cary, NC: SAS Institute Inc.