
RANDOM (CAUCHY, DF=スカラー, DRAW= 系列あるいは行列, EDF= 系列,
 EXPON, GAMMA, LAMBDA=スカラー, LAPLACE, MEAN= スカラー, NEGBIN, POISSON,
 SEEDIN= スカラー, SEEDOUT=スカラー, STDEV=スカラー, T, UNIFORM,
 VCOV= 分散共分散行列, VMEAN= 平均ベクトル)
 系列 あるいは 行列 あるいは 系列リスト ;

機能:

RANDOM は疑似乱数を発生します。正規、ポアソン、一様、負二項、ラプラス、 t 、コーシー、指数、ガンマあるいは経験分布に従う乱数を生成します。ユーザーは、場合によっては分布のパラメータを与えます。

使用法:

オプションのない RANDOM は平均がゼロで分散が1の正規乱数を (現在の SMPL の範囲のなかで) 系列として発生し、保存します。もし正規化されていない乱数が必要なときは、MEAN と STDEV オプションを含めます。

その他のオプションで、ポアソン、一様、負二項、ラプラス (二重指数)、 t 、コーシー、指数、ガンマあるいは経験分布に従う乱数を生成します。他の分布から乱数を得る方法については、例の部分参照して下さい。

TSP は現在の時間を基に毎回与える初期値を「ランダム化」しますから、毎回の結果を再現するためには固定した SEED 値を用いる必要があります。ある特定の実行に対して最初の SEED 値を変えたい場合や、それを後で使うために保存したい場合は、SEEDIN や SEEDOUT オプションを使います。

多変量正規乱数には、VCOV オプションが必要です。この (対称、あるいは直交) 行列の次元は作成する系列数になります。もしセミコロンの前にただ一つだけしか引数を書かなければ、系列はこの名前の行列で保存されます。VMEAN オプション (スカラーの MEAN オプションでなく) は非ゼロ平均ベクトルが欲しいときに用います。

経験的分布関数から乱数を発生するには、分布関数によって作られる系列 (通常は一組の残差) を与えなければなりません。この機能はブートストラップ標準誤差の計算に便利で、モデルから発生された一組の残差は DRAW オプションの入力として用いられ、以前のと同じ分布をする新しい系列は、残差の各々観測された値上に確率密度が観測値の個数分の1に等しい離散分布から観測値を抽出することによって得られます。この残差の新しい標本は、これらの乱数の関数の推定値を得るためにその後の計算で用いることが出来ます。

オプション:

CAUCHY/NOCAUCHY は発生される乱数が、コーシー分布 ($f(x) = 1/[\pi(1+x^2)]$) に従うことを示します。

DF=ステューデントの t 分布の自由度 (t オプション参照)。

DRAW= 乱数を発生するのに確率 $1/@NOB$ で標本抽出される系列名。この系列はどんな順序にもソートされていない。DRAW=行列の場合は、乱数の多変量セットが抽出されます。確率変数の数は、行列の列の数に等しい。抽出する行列は、観測値の数に等しい行の数を持たないことに注意して下さい。これは、シミュレーションやブートストラップ標準誤差には大変便利です。

EDF=経験的分布関数. DRAW=と同じ.

EXPON/NOEXPON は発生する乱数が指数分布 ($f(x) = \lambda \exp(-\lambda x)$, ここで x の期待値は $1/\lambda$, 分散は $1/\lambda^2$) に従うことを示します.

GAMMA/NOGAMMA は発生する乱数がガンマ分布 ($f(x) = (\lambda^\alpha / \text{gamfn}(\alpha)) x^{\alpha-1} \exp(-\lambda x)$, ここで x の期待値は α/λ , 分散は α/λ^2) に従うことを示します. λ を指定するのに LAMBDA=は使用しません. そのかわり MEAN=と STDEV=オプションを使います. MEANと STDEVは必ず非負です.

LAMBDA= 指数あるいは2重指数分布のパラメータ (EXPONと LAPLACE オプション参照).

LAPLACE/NOLAPLACE は発生される乱数がラプラス (2重指数) 分布 ($f(x) = \lambda \exp(-\lambda x)$, ここで $f(x) = \lambda \exp(-2|\lambda x|)$, 分散は $1/(2\lambda^2)$) に従うことを示します.

MEAN= 乱数の期待値. このオプションは正規乱数, ガンマ乱数, ポアソン乱数, そして負二項乱数にだけ適用されます. 既定値は1ですが, これはガンマと負二項乱数には向いていません.

NEGBIN/NONEGBIN は発生する乱数が負二項分布 (N,p) で, 各試行で成功確率 p の時の N 回成功 (試行回数 $-N$) するための超過待時間になります. N は整数でなくとも構いません. 平均 $= N * (1 - p)/p$, 分散 $= N * (1 - p)/p^2 = \text{平均}/p$. TSP ユーザース・ガイドの例 9.6.12 は, 別の方法を用いています: $N = \lambda$ と $p = \delta/(1 + \delta)$, あるいは $\delta = p/(1 - p)$, $\delta \geq 0$. この方法では, 平均は δ/λ , そして分散は $\lambda * (1 + \delta)/\delta^2 = \text{Mean} * (1 + \delta)/\delta$. λ を与えるのに LAMBDA= を用いません. その代わりに MEAN= と STDEV= オプションを用います. MEANは必ず非負で, STDEVは必ず平均の平方根に等しいかそれ以上でなくてはなりません.

POISSON/NOPOISSON は発生される乱数がポアソン分布 ($f(x) = \exp(-a)a^x/x!$) に従うことを示します. この分布は, 期待値と分散の両方を表す1個のパラメータ a を持ちます. a に MEAN=値で与えます.

SEEDIN= (確率SEEDの現在の値に代えて) 乱数発生を始めるための確率SEEDの値です. これは [1 と 21 億] の範囲の整数でなければいけません (それ以外の場合は, この範囲に移します). スカラーの値は倍精度で保存されますから, 有効数字が16桁まで許されます. 結果を再現したければ, 大きすぎるSEED値を与えないようにします.

SEEDOUT= 乱数発生の確率SEED (ある確率変数がこのコマンドによって作られる以前の現在の乱数SEED値). SEED は倍精度の値ですから, ここでは行列要素とするのが最もよいでしょう.

STDEV= 乱数の標準偏差. このオプションは正規乱数にだけ用います. 既定値は1になります.

T/NOT は発生される乱数が自由度 $DF=n$ のステューデント t 分布に従うことを示します.

UNIFORM/NOUNIFORM は生成される乱数がゼロと1の間の一様分布にしたがうことを示します.

VCOV= 多変量正規乱数に対する (対称な) 分散・共分散行列. (三角化した) この行列の平方根を与えることもできます.

VMEAN= 多変量正規乱数に対する平均ベクトル (既定値はゼロのベクトル).

例::

次の例は各々 1000 個の乱数を発生し, それらを与えた系列名で保存します.

```
SMPL 1 1000 ;
RANDOM STDNORM ;
RANDOM (UNIFORM) FLAT ;
FLATAB = FLAT*(B-A) + A ;      ? 区間 (A,B) で一様乱数の発生
```

RANDOM

```
CDF(CHISQ,DF=3,INV) FLAT CHI3; ? Chi-square(3)
T1EV = -LOG(-LOG(FLAT));      ? Type I Extream Value
RANDOM(CAUCHY) FAT;
RANDOM(EXPON,LAMBDA=2) Z;
RANDOM(LAPLACE,LAMBDA=0.5) DE;
RANDOM(T,DF=5) FATAIL;
RANDOM(GAMMA,MEAN=10,STDEV=2) GAMMA10;
RANDOM(NEGBIN,MEAN=10,STDEV=5) NEGBIN10;
RANDOM (MEAN=10,STDEV=3.1623) NORM10 ;
RANDOM (MEAN=10,POISSON) POISS10 ;
```

最後の2つの例は同じ平均と分散ですが、別の分布の乱数になります。平均10で標準偏差3.1623の正規変量は、次のようにしても作ることができます。

```
GENR NORM10 = 10 + STDNORM*3.1623 ;
```

次の例は相関が.5の2変量正規乱数ベクトルを作ります。

```
LOAD (TYPE=SYM,NROW=2) COVMAT ; 1 .5 1 ;
RANDOM (VCOV=COVMAT) NU1 NU2 ;
```

次に、回帰での推定された残差を用いて、同じEDFの10個の標本を作ります。

```
LIST RES RES1-RES10 ;
RANDOM (EDF=@RES) RES ;
```

終わったときにはリスト RES には元の残差系列 @RES と同じ分布を持つ10個の系列が入っています。

逆分布関数の例

```
RANDOM (UNIFORM) U ;
EV = -LOG(-LOG(U)) ;? Type I Extream Value
CDF(INV,CHISQ,DF=n) U CHIV ; ? Chi-Square (n)
```

特定の推定量の潜在的な収束率を調べるために、経験分布関数からデータを抽出するのにDRAWの行列バージョンを使った例

```
? original data set has 100 observations, compute residuals
SMPL 1 100;
OLSQ Y C X;
UNMAKE @COEF A B;
MMAKE EDF @RES X;
? estimation using 1,000 obs drawn from EDF
SMPL 1 1000 ;
RANDOM(DRAW=EDF) E X;
Y = A + B*X + E;
OLSQ Y C X;
? estimation using 10,000 obs drawn from EDF
SMPL 1 10000;
RANDOM(DRAW=EDF) E X;
Y = A + B*X + E;
OLSQ Y C X;
```

方法:

RANDOM で使われている方法は、周期 $2^{31} - 1$ で乗数 41358 の合同乗算法です。この方法は参考文献の L'Ecuyer に述べられています (そのクラスでの最適な “ランダムネス” を有しています)。この方法は整数計

算スピードと完全期間 (2^{31} 回の抽出で繰り返さない) を保証しています。いくつかのジョブ間でモンテカルロ研究を分割している場合は (毎回 TSP を再始動する), 現在の SEED を (たぶんデータバンクへ) 保存するためにジョブの最後で SEEDOUT= を使い, 次のジョブを開始するときに SEEDIN= を使います。もし TSP に現時点での開始 SEED を “ランダムイズ” させようとすれば, この乱数発生が全セットのジョブを通じて完全期間を持つことを保証しません。

合同乗算 (Multiplicative congruential) 法は, (0,1) 区間で一様な乱数を作成します。正規乱数とポアソン乱数は, 一様乱数からそれぞれ ACM アルゴリズム #488 と #369 を用いて生成されます。その他の全ての乱数はこの一様分布から, 通常漸近展開を含む逆分布関数を用いて導かれます。

参考文献:

Efron, Bradley, “Bootstrap Methods: Another Look at the Jackknife”, **Annals of Statistics**, 7, 1979, pp. 1-26.

Efron, Bradley, **The Bootstrap, the Jackknife and Other Resampling Plans**, Philadelphia: SIAM, 1982.

Efron, Bradley, and G. Gong, “A Leisurely Look at the Bootstrap, Jackknife, and Cross-validation”, **American Statistician**, February 1983, 37(1), pp. 36-48.

L’Ecuyer, Pierre, “Random Numbers for Simulation,” **Communications of the ACM**, October 1990, pp.85–97.

Schaffer, Henry E., Algorithm #369, **Collected Algorithms from ACM Volume II**, ACM, New York, N.Y..