2 確率分布

平成 28 年 5 月 8 日

概 要

PRML の「2 確率分布」について各確率分布の考察とプロットの実装

目 次

1 離散分布 **2**
1.1 ベルヌーイ分布 2
1.2 二項分布 4
1.3 幾何分布 5
1.4 超幾何分布 6
1.5 ポアソン分布 7
1.6 負の二項分布 8
1.7 ジップ分布 9
1.8 多項分布 10

2 連続分布 **11**
2.1 一様分布 11
2.2 正規分布 12
2.3 多次元正規分布 13
2.4 対数正規分布 14
2.5 ガンマ分布 15
2.6 ベータ分布 16
2.7 指数分布 17
2.8 スチューデントの t 分布 18
2.9 カイ 2 乗分布 19
2.10 F 分布 20
2.11 コーシー分布 21
2.12 アーラン分布 22
2.13 三角分布 23
2.14 ラプラス分布 24
2.15 レイリー分布 25
2.16 ロジスティック分布 26
2.17 パレート分布 27
2.18 ワイブル分布 28
2.19 フォン・ミーゼス分布 29
2.20 混合ガウス分布 30

2.21 ウィシャート分布 31
2.22 ディリクレ分布 32

1　離散分布

1.1　ベルヌーイ分布

$$Bern(x|\mu) = \mu^x(1-\mu)^{1-x} \quad , x \in \{0,1\}$$

確率変数 x は成功, 失敗を表している.
パラメータ μ は成功確率 (平均) を表している.

1.1.1　性質

二値確率変数 $x \in \{0,1\}$ について, $x=1$ となる確率 $p(x=1)$ が μ である確率変数 x 上の確率分布のこと.
例 1: コインを投げて表が出るか否かは, $Bern(x|\mu=0.5)$ に従う.
例 2: さいころを投げて 1 が出るか否かは, $Bern(x|\mu=1/6)$ に従う.
定義より

$$Bern(x|\mu) = \begin{cases} 1-\mu & (x=0) \\ \mu & (x=1) \end{cases}$$

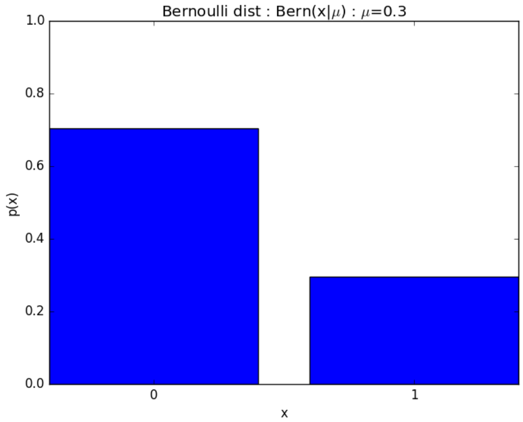
また, これより平均, 分散は

$$E[x] = \mu \quad Var[x] = \mu(1-\mu)$$

1.1.2　実装

```
mu=0.3
ONE=0
ZERO=0
for n in range(NUM):
        x=np.random.random()
        if x<mu:
                ONE+=1
        else:
                ZERO+=1
ONE/=NUM
ZERO/=NUM
X = [0,1]
Y = [ZERO,ONE]
plt.ylim(0,1)
plt.xticks(X, [0,1])
plt.title("Bernoulli␣dist␣:␣Bern(x|$\mu$)␣:␣$\mu$=%.1f" % mu)
plt.xlabel("x")
plt.ylabel("p(x)")
plt.bar(X, Y, align="center")
filename="Bernoulli.png"
plt.savefig(filename)
plt.show()
```

1.1.3 プロット



1.2 二項分布

$$Bin(x|N, \mu) = {}_N C_x \mu^x (1-\mu)^{N-x} \quad , x \in \{0, 1, \dots, N\}$$

確率変数 x は成功回数を表す.
パラメータ N と μ はそれぞれ試行回数と成功確率を表す.

1.2.1 性質

ベルヌーイ分布 $Bern(x|\mu)$ に従う試行を N 行ったときに, $x = 1$ が何回起こるかを表す.
例 1: コインを 100 回投げたときに表が出る回数は, $Bin(x|N = 100, \mu = 0.5)$ に従う.
例 2: さいころを 10 回投げた時に 1 が出る回数は, $Bin(x|N = 10, \mu = 1/6)$ に従う.
平均, 分散は

$$E[x] = N\mu \quad Var[x] = N\mu(1-\mu)$$

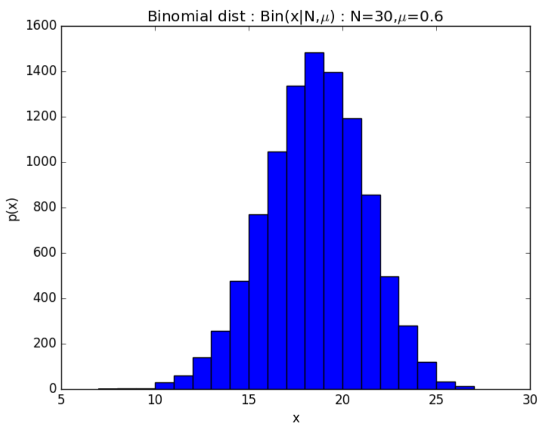
ベルヌーイ分布と関係がある.(ベルヌーイ試行)
二項分布の近似として, ガウス分布, ポアソン分布を用いることがある.

1.2.2 実装

np.random.binomial(N,μ,NUM) を用いる

```
mu,N=0.6,30
x=np.random.binomial(N,mu,NUM)
plt.title("Binomial dist : Bin(x|N,$\mu$) : N=%d,$\mu$=%.1f" % (N,mu))
plt.xlabel("x")
plt.ylabel("p(x)")
plt.hist(x,20)
filename="Binomial.png"
plt.savefig(filename)
plt.show()
```

1.2.3 プロット



1.3 幾何分布

$$Geo(x|\mu) = \mu(1-\mu)^{x-1}$$

確率変数 x は初めての成功までにかかった回数を表す.
パラメータ μ は成功確率を表す.

1.3.1 性質

ベルヌーイ分布 $Bern(x|\mu)$ を繰り返し行ったとき, 初めて $x=1$ が起こるまでの回数を表す.
例 1: コインを繰り返し投げるとき初めて表が出るまでの回数は, $Geo(x|\mu=0.5)$ に従う.
例 2: さいころを繰り返し投げるとき初めて 1 が出るまでの回数は, $Geo(x|\mu=1/6)$ に従う.
平均, 分散は

$$E[x] = 1/\mu \quad Var[x] = (1-\mu)/\mu^2$$

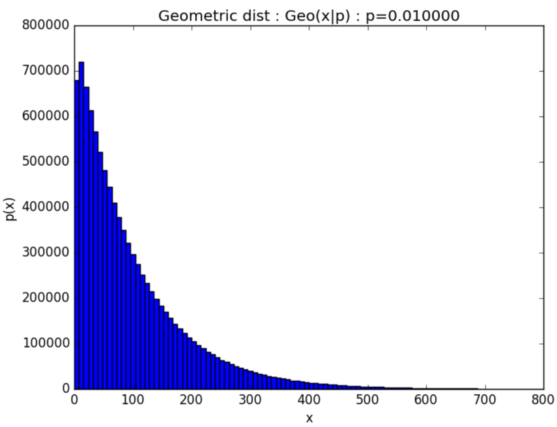
ベルヌーイ分布と関係がある.(ベルヌーイ試行)

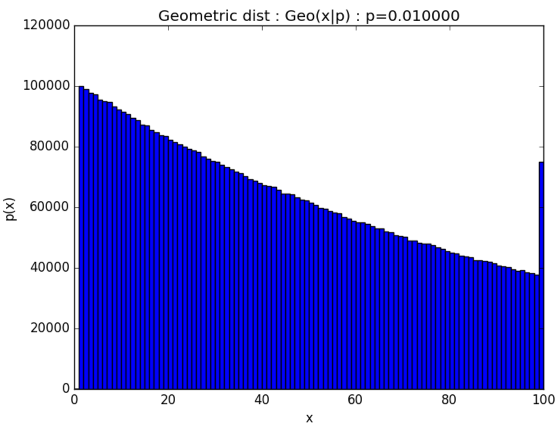
1.3.2 実装

np.random.geometric(p,NUM) を用いる.

```
p=0.01
x=np.random.geometric(p,NUM)
plt.title("Geometric␣dist␣:␣Geo(x|p)␣:␣p=%1f" %p)
plt.xlabel("x")
plt.ylabel("p(x)")
plt.hist(x,bins=100,range=(0,800))
filename="Geographic.png"
plt.savefig(filename)
plt.show()
```

1.3.3 プロット





1.4　超幾何分布

$$HGeo(x|a,b,n) = \frac{{}_aC_x \ {}_bC_{n-x}}{{}_{a+b}C_n}$$

確率変数 x は良品の数を表し.
パラメータ a, b, n でそれぞれ良品の個数, 不良品の個数, 試行回数を表す.
ただし, パラメータのとり方は一通りではなく, ここでは numpy のものに従っている.

1.4.1　性質

a 個の A と b 個の B の中から n 個とりだしたとき A を取り出せる個数を表す.(非復元抽出)
例 1: 200 個の内 190 個の良品と 10 個の不良品が入った袋から 20 個の製品をとりだしたときに含まれる良品の個数は, $Geo(x|a = 190, b = 10, n = 20)$ に従う.
平均, 分散は

$$E[x] = \frac{aN}{a+b} \quad Var[x] = \frac{Nab(a+b-N)}{(a+b)^2(a+b-1)}$$

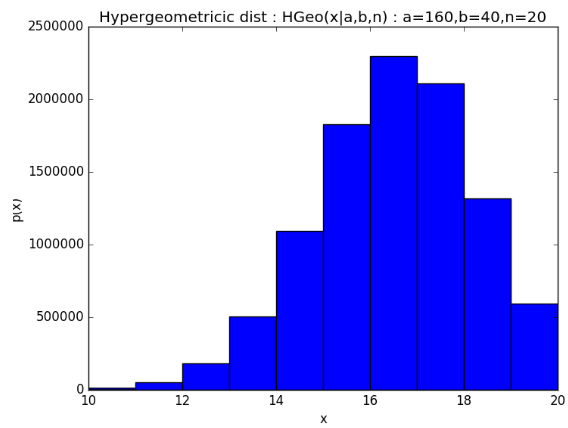
二項分布を非復元抽出に適応したようなもの.(ベルヌーイ試行ではない)
非復元抽出ではあるが, a,b ともに大きいときく n が小さいときは二項分布で近似できる.

1.4.2　実装

np.random.Hypergeometric(a,b,n,NUM) を用いる.

```
a,b,n=160,40,20
x=np.random.hypergeometric(a,b,n,NUM)
plt.title("Hypergeometricic␣dist␣:␣HGeo(x|a,b,n)␣:␣a=%d,b=%d,n=%d" %(a,b,n))
plt.xlabel("x")
plt.ylabel("p(x)")
plt.hist(x,range(10,21))
filename="Hypergeometric.png"
plt.savefig(filename)
plt.show()
```

1.4.3　プロット



1.5 ポアソン分布

$$Po(x|\lambda) = e^{-\lambda}\frac{\lambda^x}{x!} \quad x \in \{0, 1, \dots\}$$

確率変数 x は単位時間当たりの発生回数を表す.
パラメータ λ は単位時間当たりの平均発生回数を表す.

1.5.1 性質

所与の時間間隔で発生する離散的な事象を数える特定の確率変数 X を持つ離散確率分布のこと.
例 1: 平均 10 回/1 時間車が通る道に 1 時間で通る車の台数は, $Po(x|\lambda = 10)$ に従う.
平均, 分散は

$$E[x] = \lambda \quad Var[x] = \lambda$$

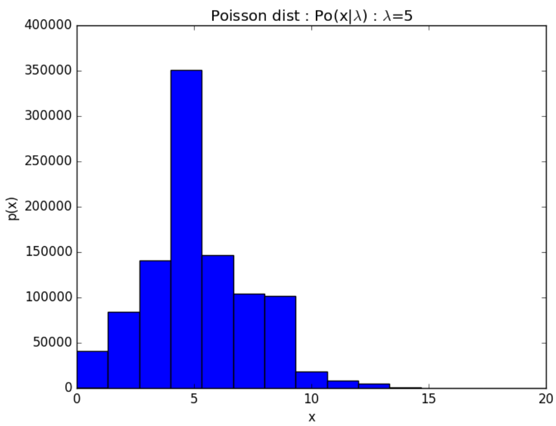
めったに起こらないことに対して用いることが多い.

1.5.2 実装

np.random.poisson(lam,N) を用いる.

```
lam=5
x=np.random.poisson(lam,NUM)
plt.title("Poisson␣dist␣:␣Po(x|$\lambda$)␣:␣$\lambda$=%d" % lam)
plt.xlabel("x")
plt.ylabel("p(x)")
plt.hist(x,15)
filename="Poisson.png"
plt.savefig(filename)
plt.show()
```

1.5.3 プロット



1.6 負の二項分布

$$p(N|n,p) = {}_{N+n-1}C_{x\ n-1}p^n(1-p)^N$$

確率変数 N は失敗回数を表す.
パラメータ n, p は n 回目の成功, 成功確率を表す. ただし, パラメータのとり方は複数あるが numpy のものに従うとする.

1.6.1 性質

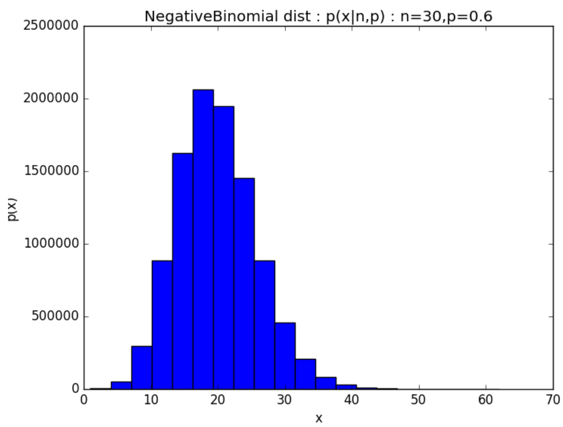
$Bin(x|p)$ に従うベルヌーイ試行を続けて行ったとき n 回目の成功までに N 回失敗する確率を表す.

1.6.2 実装

np.random.negative_binomial(n,p,NUM) を用いる.

```
p=0.6
n=30
x=np.random.negative_binomial(N,mu,NUM)
plt.title("NegativeBinomial␣dist␣:␣p(x|n,p)␣:␣n=%d,p=%.1f" % (n,p))
plt.xlabel("x")
plt.ylabel("p(x)")
plt.hist(x,20)
filename="NegativeBinomial.png"
plt.savefig(filename)
plt.show()
```

1.6.3 プロット



1.7 ジップ分布

$$f(x|s,N) = \frac{1/x^s}{\sum_{n=1}^{N} 1/n^s}$$

確率変数 x は順位.
パラメータ s は下のジップ則に従うなら常に $s = 1$, N は全体数.
この分布は, 全体数 N のとき x 番目の占有率を表すと考えてよい.

1.7.1 性質

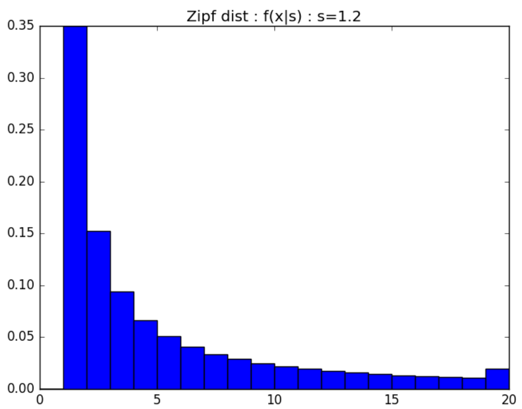
ジップの法則とは, 出現頻度が k 番目に大きい要素が全体に占める割合が $1/k$ に比例するという経験則である. また, この法則が機能する世界を「ジフ構造」と記する論者もいる.
包括的な理論的説明はまだ成功していないものの, 様々な現象に適用できることが知られている.
この法則に従う確率分布（離散分布）をジップ分布という.
ジップ分布はゼータ分布の特殊な形である. パレート分布の離散版.
ウェブのアクセス頻度, 都市の人口などのモデル化に用いられる.

1.7.2 実装

np.random.Zipf(s,NUM) を用いる.

```
s=1.2
x=rd.zipf(s,NUM)
plt.hist(x,bins=20, range=(0,20), normed=True)
plt.title("Zipf␣dist␣:␣f(x|s)␣:␣s=%s" % s)
filename="Zipf.png"
plt.savefig(filename)
plt.show()
```

1.7.3 プロット



1.8 多項分布

$$Mult(x|\mu, N) = \frac{N!}{x_1! \dots x_K} \Pi_{k=1}^{K} \mu_k^{x_k}$$

確率変数 x は成功回数を表す.
パラメータ μ は成功確率, N は試行回数を表す.

1.8.1 性質

二項分布の拡張. 独立な二項分布を K 個集めたようなもの.
平均, 共分散は

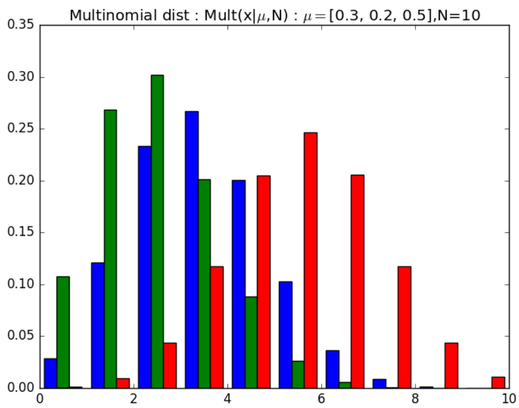
$$E[x_i] = N\mu_i \quad var[x_i] = N\mu_i(1-\mu_i) \quad cov[x_i, x_j] = -N\mu_i\mu_j$$

1.8.2 実装

np.random.multinomial(N,mu,NUM) を用いる.

```
N,mu=10,[0.3,0.2,0.5]
x=rd.multinomial(N,mu,NUM)
plt.hist(x, normed=True)
plt.title("Multinomial␣dist␣:␣Mult(x|$\mu$,N)␣:␣$\mu=$%s,N=%s" %(mu,N))
filename="Multinomial.png"
plt.savefig(filename)
plt.show()
```

1.8.3 プロット



2　連続分布

2.1　一様分布

$$U(x|a,b) = \frac{1}{b-a} \quad ,x \in [a,b]$$

パラメータは, a,b で区間を表す.

2.1.1　性質

区間の中の値を一定の確率でとるという分布.
平均, 分散は

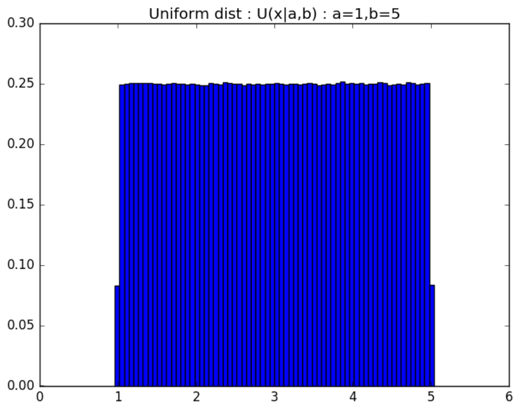
$$E[x] = \frac{a+b}{2} \quad Var[x] = \frac{(b-a)^2}{12}$$

2.1.2　実装

np.random.uniform(a,b,NUM) を用いる.

```
a,b=1,5
x=rd.uniform(a,b,NUM)
plt.hist(x, bins=100, range=(0, 6), normed=True)
plt.title("Uniform dist : U(x|a,b) : a=1,b=5")
filename="Uniform.png"
plt.savefig(filename)
plt.show()
```

2.1.3　プロット



2.2 正規分布

$$N(x|\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\} \quad ,x \in R$$

パラメータは, μ と σ でそれぞれ平均と分散を表す.

2.2.1 性質

最もエントロピーの大きな分布.
平均, 分散は

$$E[x] = \mu \quad Var[x] = \sigma^2$$

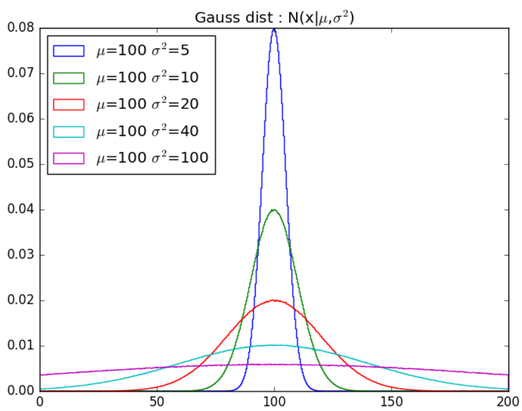
例 1: 実験における測定の誤差は正規分布に従って分布すると仮定さる.
実際のモデル化においては R 全体をとるものは少ないため加工して用いる.
ガウス分布の平均に対する共役事前分布としてはガウス分布が用いられる.

2.2.2 実装

np.random.normal(mu,sig,NUM) を用いる.

```
mu=100
for sig in [5,10,20,40,100]:
        x=rd.normal(mu,sig,NUM)
        label = "$\mu$=%s␣$\sigma^2$=%s" %(mu,sig)
        plt.hist(x, histtype='step', bins=500, range=(0, 200),
                normed=True, label=label)
plt.title("Gauss␣dist␣:␣N(x|$\mu$,$\sigma^2$)")
plt.legend(loc="upper␣left")
filename="gauss.png"
plt.savefig(filename)
plt.show()
```

2.2.3 プロット



2.3 多次元正規分布

$$N(x|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} exp\{-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)\} \quad ,x \in R^D$$

パラメータは, μ と Σ でそれぞれ平均と共分散を表す.

2.3.1 性質

ガウス分布の多次元版.
平均, 共分散は

$$E[x] = \mu \quad cov[x] = \Sigma$$

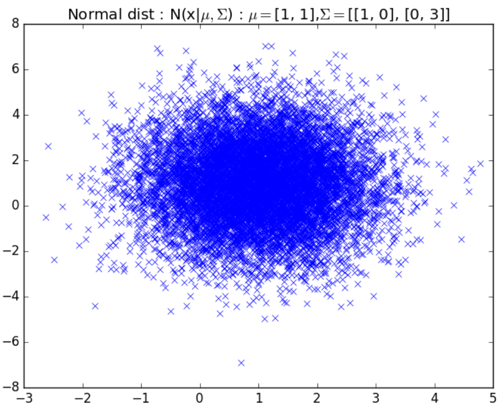
平均は分布の場所, 共分散行列は分布の形を決める.
多次元ガウス分布の平均についての共役事前分布には多次元ガウス分布を用いる.

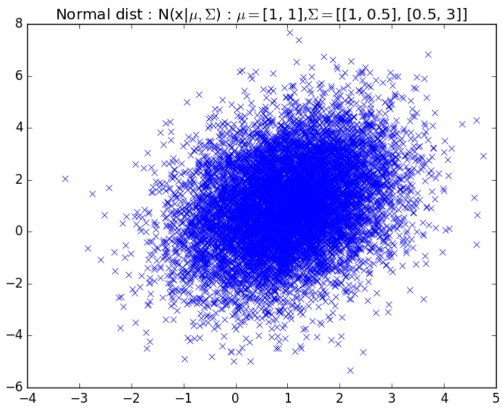
2.3.2 実装

np.random.malutivariate_normal(mu,cov,NUM) を用いる.

```
mu=[1,1]
cov=[[1,0.5],[0.5,3]]
x,y=rd.multivariate_normal(mu,cov,NUM).T
plt.plot(x,y,'x')
plt.title("Normal␣dist␣:␣N(x|$\mu,\Sigma$)␣:␣$\mu=$%s,$\Sigma=$%s" %(mu,cov))
filename="Normal_N_2.png"
plt.savefig(filename)
plt.show()
```

2.3.3 プロット





2.4　対数正規分布

$$N(x|\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} exp\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\} \quad ,x>0$$

ただし, パラメータは, μ と σ でそれぞれ平均と分散を表さない.

2.4.1　性質

対数正規分布に従う確率変数 x に対して, 新たに確率変数 $y = \ln x$ をとったとき, この確率変数 y が正規分布に従う.
平均, 分散は

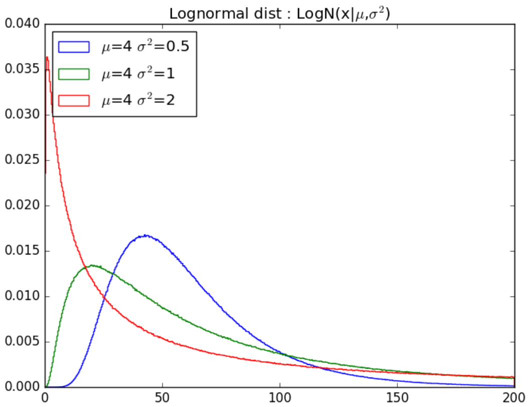
$$E[x] = e^{\mu+\sigma^2} \quad Var[x] = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

2.4.2　実装

np.random.lognormal(mu,sig,NUM) を用いる.

```
mu=4
for sig in [0.5,1,2]:
        x=rd.lognormal(mu,sig,NUM)
        label = "$\mu$=%s $\sigma^2$=%s" %(mu,sig)
        plt.hist(x, histtype='step', bins=500, range=(0, 200),
                normed=True, label=label)
plt.title("Lognormal dist : LogN(x|$\mu$,$\sigma^2$)")
plt.legend(loc="upper left")
filename="Lognormal.png"
plt.savefig(filename)
plt.show()
```

2.4.3　プロット



2.5 ガンマ分布

$$Gam(x|a,b) = \frac{1}{\Gamma(a)} b^a x^{a-1} exp(-bx) \quad ,x \geq 0$$

パラメータは, a と b で

2.5.1 性質

平均, 分布は

$$E[x] = \frac{a}{b} \quad Var[x] = \frac{a}{b^2}$$

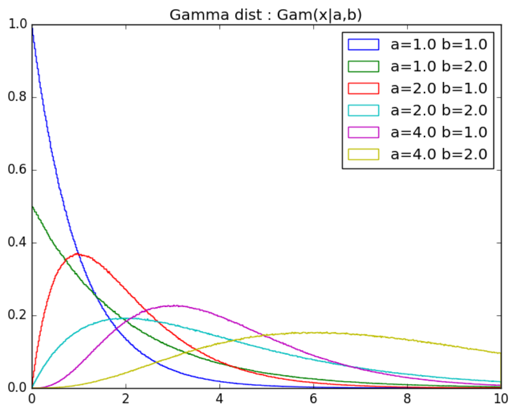
ガウス分布の精度に対する共役事前分布として用いられる.

2.5.2 実装

np.random.gamma(a,b,NUM) を用いる.

```
for a in [1.0,2.0,4.0]:
    for b in [1.0,2.0]:
        data = list()
        for i in range(NUM):
            data.append(rd.gamma(a, b))
        label = "a=%s␣b=%s" %(a, b)
        plt.hist(np.array(data), histtype='step', bins=500, range=(0, 10),
                 normed=True, label=label)
plt.title("Gamma␣dist␣:␣Gam(x|a,b)")
plt.legend(loc="upper␣right")
filename="Gamma.png"
plt.savefig(filename)
plt.show()
```

2.5.3 プロット



2.6 ベータ分布

$$Beta(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad , x \in [0,1]$$

また, a,b はパラメータである.

2.6.1 性質

平均, 分散は

$$E[x] = \frac{a}{a+b} \quad Var[x] = \frac{ab}{(a+b)^2(a+b+1)}$$

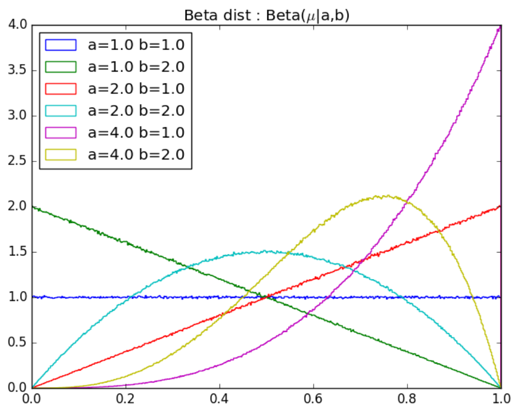
ベルヌーイ分布, 二項分布の成功確率に対する共役事前分布に用いられる.

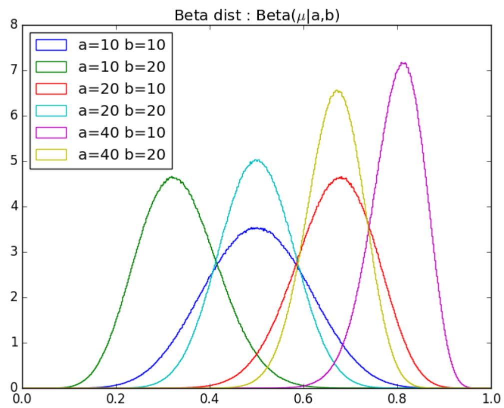
2.6.2 実装

np.random.beta(a,b,NUM) を用いる.

```
for a in [10,20,40]:
    for b in [10,20]:
        data = list()
        for i in range(NUM):
            data.append(rd.beta(a, b))
        label = "a=%s␣b=%s" %(a, b)
        plt.hist(np.array(data), histtype='step', bins=500, range=(0.0, 1.0),
                    normed=True, label=label)
plt.title("Beta␣dist␣:␣Beta($\mu$|a,b)")
plt.legend(loc="upper␣left")
filename="Beta_2.png"
plt.savefig(filename)
plt.show()
```

2.6.3 プロット





2.7 指数分布

$$
Ex(x|\lambda) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}
$$

パラメータは,

2.7.1 性質

ポアソン分布が単位時間の生起確率を示し, 指数分布は生起期間の確率を示す.
平均, 分散は

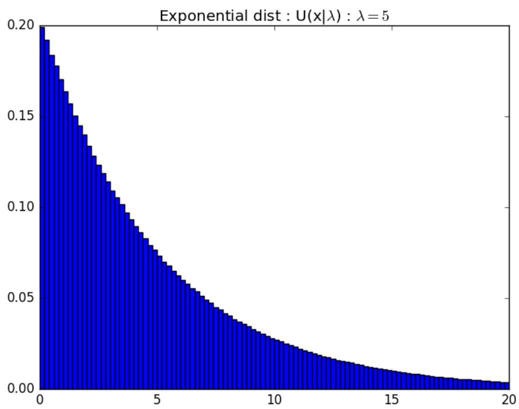
$$
E[x] = \frac{1}{\lambda} \;\; Var[x] = \frac{1}{\lambda^2}
$$

2.7.2 実装

np.random.exponential(lam,NUM) を用いる.

```
lam=5
x=rd.exponential(lam,NUM)
plt.hist(x, bins=100, range=(0, 20), normed=True)
plt.title("Exponential dist : Ex(x|$\lambda$) : $\lambda=5$")
filename="Exponential.png"
plt.savefig(filename)
plt.show()
```

2.7.3 プロット



2.8 スチューデントの t 分布

$$St(x|\mu,\lambda,\nu) = \frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)}\left(\frac{\lambda}{\pi\nu}\right)^{1/2}\left[1+\frac{\lambda(x-\mu)^2}{\nu}\right]^{-\nu/2-1/2} \quad ,x \in R$$

2.8.1 性質

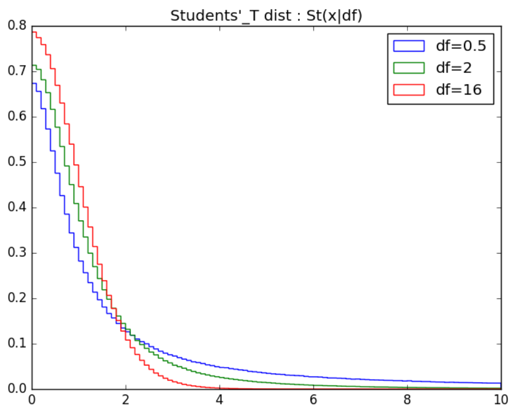
ガウス分布の精度パラメータの共役事前分布としてガンマ分布を用いたときの事後分布. ガウス分布よりも分布の裾が長く, それゆえ頑健性を持つとされる.

2.8.2 実装

np.random.standard_t(df,NUM) を用いる.

```
for df in [0.5,2,16]:
        x=rd.standard_t(df,NUM)
        label = "df=%s" %df
        plt.hist(x, histtype='step', bins=500, range=(0, 10),
                normed=True, label=label)
plt.title("Students'_T␣dist␣:␣St(x|df)")
plt.legend(loc="upper␣right")
filename="T.png"
plt.savefig(filename)
plt.show()
```

2.8.3 プロット



2.9 カイ 2 乗分布

$$p(x) = \frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$

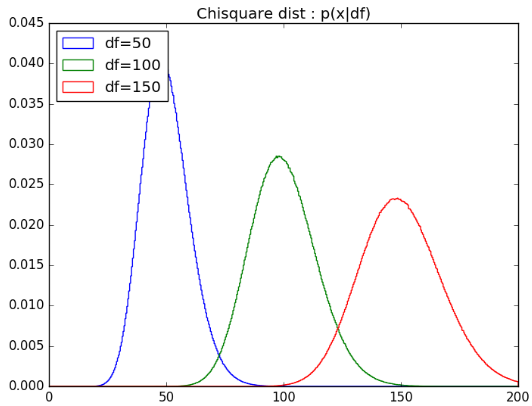
2.9.1 性質

2.9.2 実装

np.random.chisquare(df,NUM) を用いる.

```
for df in [50,100,150]:
        x=rd.chisquare(df,NUM)
        label = "df=%s" % df
        plt.hist(x, histtype='step', bins=500, range=(0, 200),
                normed=True, label=label)
plt.title("Chisquare␣dist␣:␣p(x|df)")
plt.legend(loc="upper␣left")
filename="Chisquare.png"
plt.savefig(filename)
plt.show()
```

2.9.3 プロット



2.10 F 分布

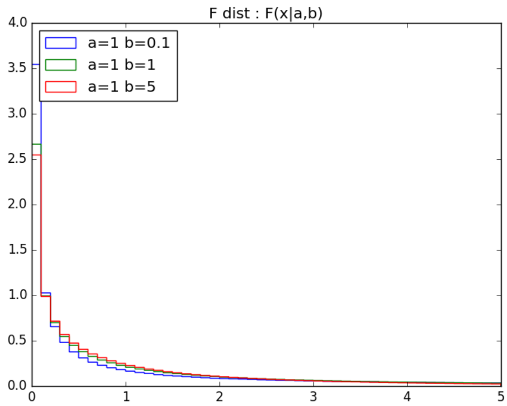
2.10.1 性質

2.10.2 実装

np.random.chisquare(df,NUM) を用いる.

```
for a in [1]:
        for b in [0.1,1,5]:
                x=rd.f(a,b,NUM)
                label = "a=%s␣b=%s" %(a,b)
                plt.hist(x, histtype='step', bins=50, range=(0,5),
                      normed=True, label=label)
plt.title("F␣dist␣:␣F(x|a,b)")
plt.legend(loc="upper␣left")
filename="F.png"
plt.savefig(filename)
plt.show()
```

2.10.3 プロット



2.11　コーシー分布

$$p(x|x_0, \gamma) = \frac{1}{\pi\gamma[1 + \left(\frac{x - x_0}{\gamma}\right)^2]}$$

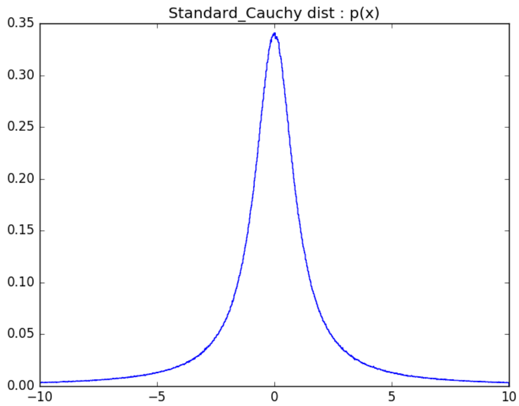
2.11.1　性質

2.11.2　実装

np.random.standard_cauchy(NUM) を用いる.

```
x=rd.standard_cauchy(NUM)
plt.hist(x, histtype='step', bins=100, range=(0, 10), normed=True)
plt.title("Standard_Cauchy␣dist␣:␣p(x)")
plt.legend(loc="upper␣right")
filename="cauchy.png"
plt.savefig(filename)
plt.show()
```

2.11.3　プロット



2.12 アーラン分布

2.13 三角分布

$$p(x|l,m,r) = \begin{cases} \frac{2(x-l)}{(r-l)(m-l)} & (l \le x \le m) \\ \frac{2(m-x)}{(r-l)(r-m)} & (m \le x \le r) \\ 0 & (otherwise) \end{cases}$$

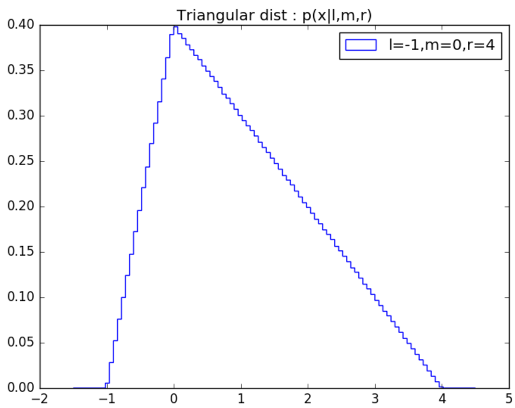
2.13.1 性質

2.13.2 実装

np.random.triangular(l,m,r,NUM) を用いる.

```
x=rd.triangular(l,m,r,NUM)
label = "l=%s,m=%s,r=%s" % (l,m,r)
plt.hist(x, histtype='step', bins=100, range=(-1.5, 4.5),
          normed=True, label=label)
plt.title("Triangular␣dist␣:␣p(x|l,m,r)")
plt.legend(loc="upper␣right")
filename="Triangular.png"
plt.savefig(filename)
plt.show()
```

2.13.3 プロット



2.14 ラプラス分布

$$p(x|\mu,\lambda) = \frac{1}{2\lambda} exp\left(-\frac{|x-\mu|}{\lambda}\right)$$

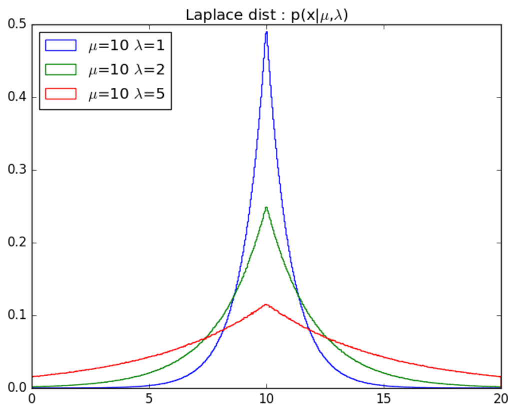
2.14.1 性質

2.14.2 実装

np.random.laplace(mu,lam,NUM) を用いる.

```
for mu in [10]:
        for lam in [1,2,5]:
                x=rd.laplace(mu,lam,NUM)
                label = "$\mu$=%s␣$\lambda$=%s" %(mu,lam)
                plt.hist(x, histtype='step', bins=500, range=(0,20),
                      normed=True, label=label)
plt.title("Laplace␣dist␣:␣p(x|$\mu$,$\lambda$)")
plt.legend(loc="upper␣left")
filename="Laplace.png"
plt.savefig(filename)
plt.show()
```

2.14.3 プロット



2.15 レイリー分布

$$p(x|s) = \frac{x}{s^2} e^{-\frac{x^2}{2s^2}}$$

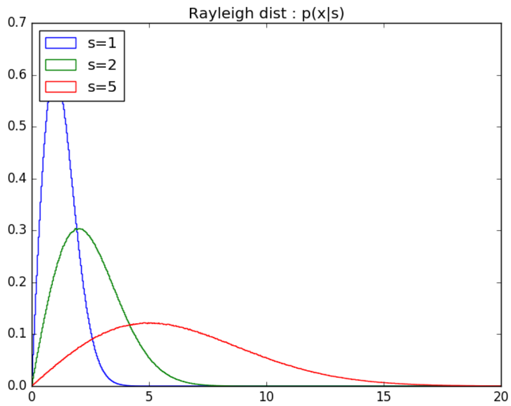
2.15.1 性質

2.15.2 実装

np.random.rayleigh(s,NUM) を用いる.

```
for s in [1,2,5]:
        x=rd.rayleigh(s,NUM)
        label = "s=%s" % s
        plt.hist(x, histtype='step', bins=500, range=(0,20),
                normed=True, label=label)
plt.title("Rayleigh␣dist␣:␣p(x|s)")
plt.legend(loc="upper␣left")
filename="Rayleigh.png"
plt.savefig(filename)
plt.show()
```

2.15.3 プロット



2.16 ロジスティック分布

$$p(x|\mu,\lambda) = \frac{e^{-(x-\mu)/s}}{s(1+e^{-(x-\mu)/s})^2}$$

パラメータは, l,m,r で三角形の左端, 頂点, 右端の点を表す.

2.16.1 性質

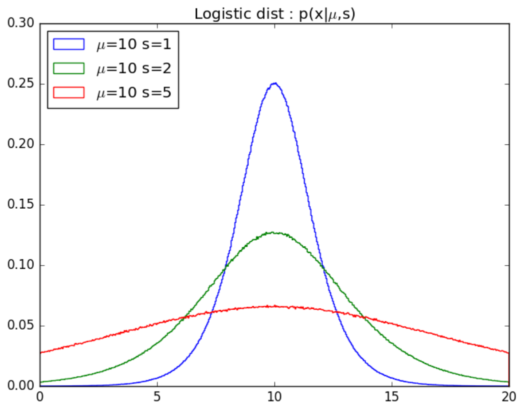
ベータ分布の代用として用いられることがある.

2.16.2 実装

np.random.logistic(mu,s,NUM) を用いる.

```
for mu in [10]:
        for s in [1,2,5]:
                x=rd.logistic(mu,s,NUM)
                label = "$\mu$=%s␣s=%s" %(mu,s)
                plt.hist(x, histtype='step', bins=500, range=(0,20),
                      normed=True, label=label)
plt.title("Logistic␣dist␣:␣p(x|$\mu$,s)")
plt.legend(loc="upper␣left")
filename="Logistic.png"
plt.savefig(filename)
plt.show()
```

2.16.3 プロット



2.17 パレート分布

$$p(x) = \frac{am^a}{x^{a+1}}$$

2.17.1 性質

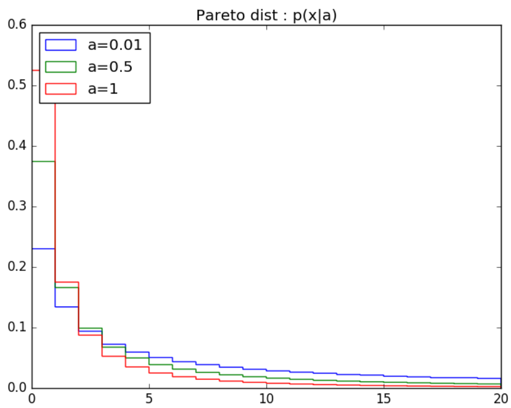
スケールフリーという性質を持つ.
所得をモデル化する際に用いられたりする.

2.17.2 実装

np.random.pareto(a,NUM) を使う.

```
for a in [0.01,0.5,1]:
        x=rd.pareto(a,NUM)
        label = "a=%s" % a
        plt.hist(x, histtype='step', bins=20, range=(0, 20),
                normed=True, label=label)
plt.title("Pareto␣dist␣:␣p(x|a)")
plt.legend(loc="upper␣left")
filename="Pareto.png"
plt.savefig(filename)
plt.show()
```

2.17.3 プロット



2.18 ワイブル分布

$$p(x) = \frac{a}{\lambda}\left(\frac{x}{\lambda}\right)^{a-1} e^{-(x/\lambda)^a}$$

2.18.1 性質

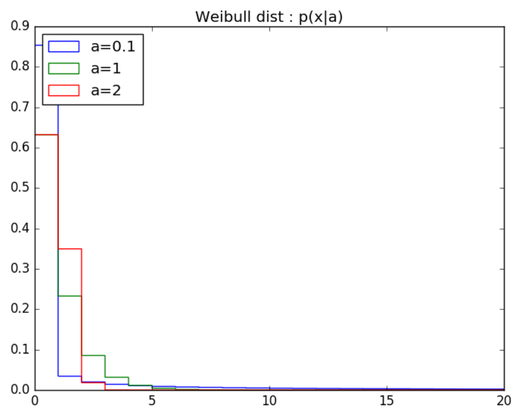
時間に対する劣化や寿命を統計的に記述するためにも用いられる.

2.18.2 実装

np.random.Weibull(a,NUM) を使う.

```
for a in [0.1,1,2]:
        x=rd.weibull(a,NUM)
        label = "a=%s" % a
        plt.hist(x, histtype='step', bins=20, range=(0, 20),
                normed=True, label=label)
plt.title("Weibull␣dist␣:␣p(x|a)")
plt.legend(loc="upper␣left")
filename="Weibull.png"
plt.savefig(filename)
plt.show()
```

2.18.3 プロット



2.19 フォン・ミーゼス分布

$$p(\theta|\theta_0, m) = \frac{1}{2\pi I_0(m)} exp\{mcos(\theta - \theta_0)\} \quad \theta in[-\pi, \pi)$$

パラメータ θ_0 は平均, m は集中度パラメータである.

2.19.1 性質

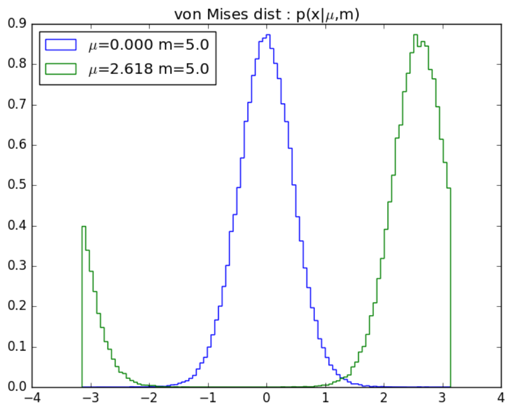
周期的なガウス分布に従う確率変数に対応する.

2.19.2 実装

np.random.vonmises(mu,m,NUM) を使う.

```
m = 5.0
for mu in [0,5*pi/6]:
        label="$\mu$=%.3f m=%s" %(mu,m)
        x = np.random.vonmises(mu, m, 100000)
        plt.hist(x,histtype="step", bins=100, range=(-pi,pi),
                normed=True,label=label)
plt.title("von Mises dist : p(x|$\mu$,m)")
plt.legend(loc="upper left")
filename="Von_Mises.png"
plt.savefig(filename)
plt.show()
```

2.19.3 プロット



2.20 混合ガウス分布

$$p(x|\pi, \mu\sigma^2) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \sigma_k^2)$$

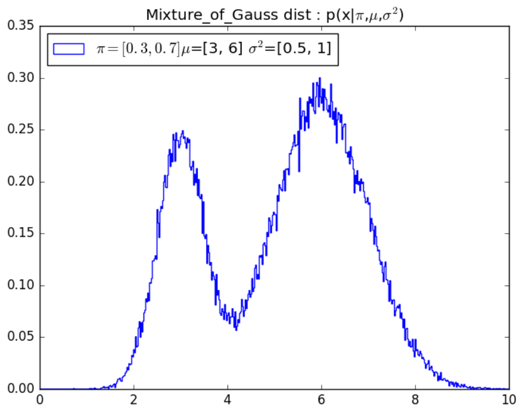
2.20.1 性質

2.20.2 実装

```
p=0.3
mu,sig=[3,6],[0.5,1]
x=np.array([])
for n in range(NUM):
        y=np.random.rand()
        if y<p:
                x=np.append(x,rd.normal(mu[0],sig[0]))
        else:
                x=np.append(x,rd.normal(mu[1],sig[1]))

label = "$\pi=[0.3,0.7]␣\mu$=%s␣$\sigma^2$=%s" %(mu,sig)
plt.hist(x, histtype='step', bins=500, range=(0,10), normed=True, label=label)
plt.title("Mixture_of_Gauss␣dist␣:␣p(x|$\pi$,$\mu$,$\sigma^2$)")
plt.legend(loc="upper␣left")
filename="MixtureGauss.png"
plt.savefig(filename)
plt.show()
```

2.20.3 プロット



2.21 ウィシャート分布

2.21.1 性質

多次元正規分布のパラメータ Σ の共役事前分布として用いられる.

2.21.2 実装

sp.stats.wishert(df,Σ,NUM) を用いる.

2.22 ディリクレ分布

$$Dir(x|a) = \frac{\Gamma(a_0)}{\Gamma(a_1)\dots\Gamma(a_K)}\Pi_{k=1}^{K}x_k^{a_k-1}$$

パラメータは a 平均ベクトルである.

2.22.1 性質

多項分布のパラメータ μ の共役事前分布として用いられる.

2.22.2 実装

np.dirichret(a,NUM) を用いる.

```
a=[0.5,2,4]
x=rd.dirichlet(a,NUM)
plt.hist(x,histtype='step', bins=100, range=(0, 1), normed=True)
plt.title("Dirichlet␣dist␣:␣Dir(x|a)␣:␣a%s" %a)
filename="Dirichlet.png"
plt.savefig(filename)
plt.show()
```

2.22.3 プロット