《天体物理中的统计方法》– 第 3 次作业 *

苏镇波†

中国科学技术大学物理学院天文系，合肥 230026

Gamma 分布的 mgf:

$$M(t) = \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \tag{1}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^x (t-\lambda) dx \tag{2}$$

积分在 $t < \lambda$ 的时候会发散，因此可以写成:

$$M(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} (\frac{\Gamma(\alpha)}{(\lambda - t)^\alpha}) = (\frac{\lambda}{\lambda - t})^\alpha \tag{3}$$

求导后，可得:

$$M'(0) = E(X) = \frac{\alpha}{\lambda} \tag{4}$$

$$M''(0) = E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2} \tag{5}$$

对于方差，有:

$$Var(X) = E(X^2) - [E(X)]^2 \tag{6}$$

$$= \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} \tag{7}$$

$$= \frac{\alpha}{\lambda^2} \tag{8}$$

如果 X 和 Y 分别遵循参数为 α_1, λ、α_2, λ 得 gamma 分布，则 $X + Y$ 的 mgf 为:

$$(\frac{\lambda}{\lambda - t})^{\alpha_1} (\frac{\lambda}{\lambda - t})^{\alpha_2} = (\frac{\lambda}{\lambda - t})^{\alpha_1 + \alpha_2} \tag{9}$$

即具有相同 λ 值的独立 gamma 随机变量之和遵循伽马分布，因此具有 n 自由度的 chi-square 分布为 $\alpha = n/2, \lambda = \frac{1}{2}$ 的 gamma 分布，则有:

$$f(v) = \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1} e^{-v/2} \tag{10}$$

其 mgf 为:

$$M(t) = (1 - 2t)^{-n/2} \tag{11}$$

*2022 春季《天体物理中的统计方法》

†Email: zbsu@mail.ustc.edu.cn, 学号: SA21022002

1 Problem 1

Answer.

如果 U,V 是互相独立的，且 $U \sim \chi_n^2, V \sim \chi_n^2$，那么根据式 (9)，$U+V$ 的 mgf 为:

$$(\frac{\lambda}{\lambda - t})^{n/2}(\frac{\lambda}{\lambda - t})^{m/2} = (\frac{\lambda}{\lambda - t})^{(m+n)/2} \tag{12}$$

即 $U + V \sim \chi_n^2$。

2 Problem 2

Answer.

根据式 (1-3) 的证明，对于 chi-square 而言，$\alpha = n/2$，$\lambda = \frac{1}{2}$，带入式 (3) 中，可得 χ_n^2 分布的 mgf 为:

$$M(t) = (1 - 2t)^{-n/2} \tag{13}$$

将 $\alpha = n/2, \lambda = \frac{1}{2}$ 带入式 (4) 和式 (8)，可得：平均数为 n，方差为 $2n$。

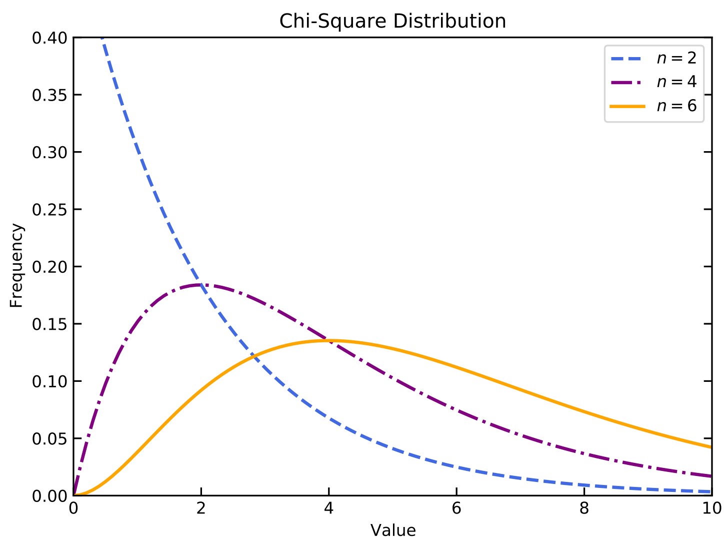

图 1: 如图，蓝色、紫色和橙色分别对应 Chi-square 下，n=2,4,6 的情况。

3 Problem 3

Answer.

$Y=\ln x$，根据 mgf 的定义，并且 $E(\mathrm{e}^{tlnx})=E(x^t)$ 有：

$$M_Y(t) = (\lambda^\alpha/\Gamma(\alpha))\int x^t x^{\alpha-1}e^{-\lambda x}dx \tag{14}$$
$$= (\lambda^\alpha/\Gamma(\alpha))\int x^{t+\alpha-1}e^{-\lambda x}dx \tag{15}$$
$$= (\lambda^\alpha/\Gamma(\alpha))\Gamma(t+\alpha)/\lambda^{t+\alpha} \tag{16}$$

求导后，可得：

$$M'(t) = -\lambda^{-\mathrm{t}}\ln\lambda \tag{17}$$
$$M''(t) = \lambda^{-\mathrm{t}}\ln^2\lambda \tag{18}$$

则有：

$$M'(0) = E(X) = -\ln\lambda \tag{19}$$
$$M''(0) = E(X^2) = \ln^2\lambda \tag{20}$$

因此均值和方差为：

$$E(X) = -\ln\lambda \tag{21}$$
$$Var(X) = E(X^2) - [E(X)]^2 \tag{22}$$
$$= \ln^2\lambda - (-\ln\lambda)^2 \tag{23}$$
$$= 0 \tag{24}$$

4 Problem 4

Answer.

考虑 $Y=X^{\frac{1}{3}}$，X 在区间 [0, 1] 上，有：

$$E(Y) = \int_0^1 x^{1/3}dx = \frac{3}{4} \tag{25}$$
$$E(Y^2) = \int_0^1 x^{2/3}dx = \frac{3}{5} \tag{26}$$
$$Var(Y) = E(Y^2) - E^2(Y) = \frac{3}{5} - \frac{3}{4}^2 = 3.75\% \tag{27}$$

X 在区间 [1, 2] 上，有：

$$E(Y) = \int_1^2 x^{1/3}dx \approx 1.13988 \tag{28}$$
$$E(Y^2) = \int_1^2 x^{2/3}dx \approx 1.3049 \tag{29}$$
$$Var(Y) = E(Y^2) - E^2(Y) = 1.3049 - 1.13988^2 = -41.08\% \tag{30}$$

$3.75\% < 41.08\%$，则在 [0, 1] 区间上效果更好。

5 Problem 5

Answer.

根据高斯分布，利用 MCMC 方法有：

$$\hat{I}(f) = \frac{1}{N}(\frac{1}{\sqrt{2\pi}})\sum_{i=1}^{N} e^{-X_i^2/2} \tag{31}$$

代码形式为：

```python
import numpy as np
import random

def mcmc(N, x_range):
    """
    N: Numbers of MCMC method iteration times.
    x_range: integration limits
    """
    sum = 0
    for i in range(N):
        x = random.uniform(0, int(x_range))
        sum += np.exp(-x**2/2)

    return 1/N * (1/np.sqrt(2*np.pi)) * sum

print(N = mcmc(10000, x_range = 1)) #0.34174123263753037

print(N = mcmc(10000, x_range = 5)) #0.09880943270473465
```

当积分区间为 [0, 1] 时，结果为 0.3417；

当积分区间为 [0, 5] 时，结果为 0.0988，不接近 $\frac{1}{2}$。

6 Problem 6

Answer.

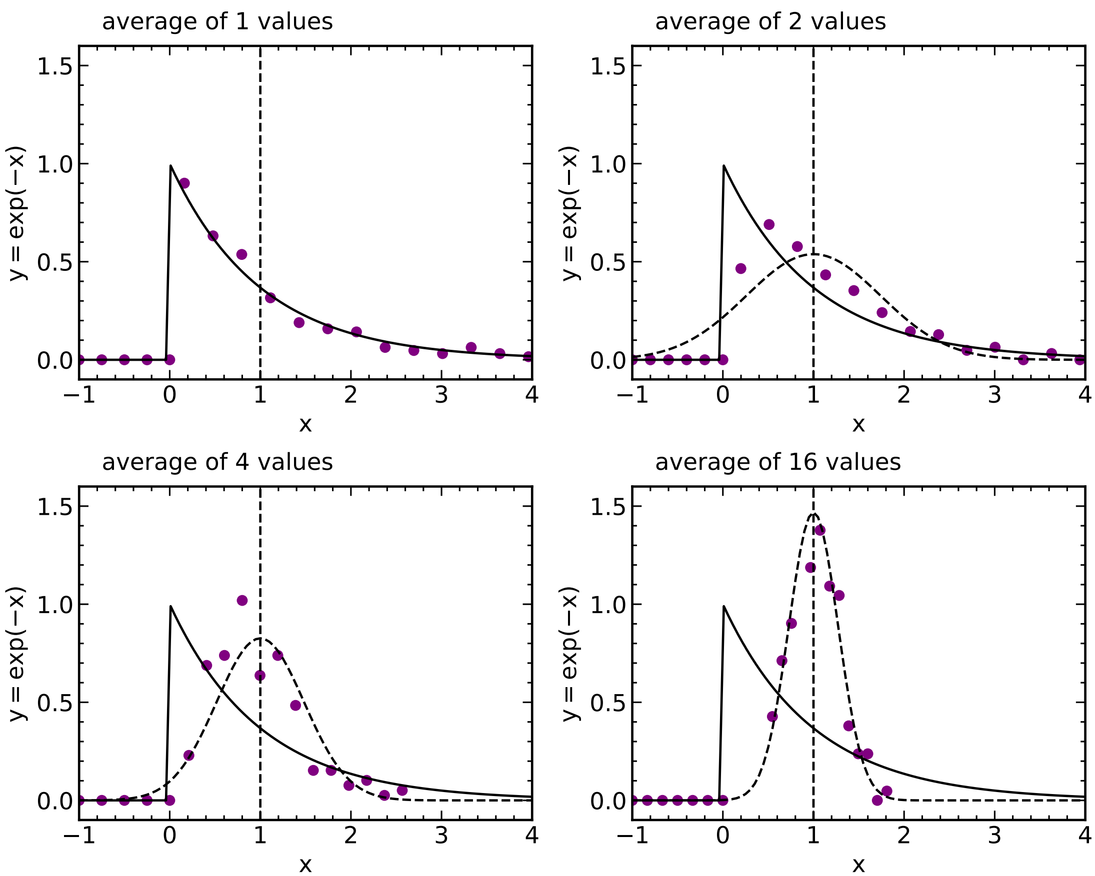

图 2: 上图中，紫色原点为随机生成的指数分布，黑色实线为初始指数分布，黑色虚线为 overplot 的最佳拟合高斯轮廓。分别为平均 1、2、4、16 个，生成随机数为 200。