2.5.1 Evaluation of Expectation: Numerical Integration

Univariate Case: Consider integration of a function $f(x)$.

Suppose that x is a scalar.

Let $x_0, x_1, x_2, \cdots, x_n$ be n nodes, which are sorted by order of size but not necessarily equal intervals between x_{i-1} and x_i for $i = 1, 2, \cdots, n$.

Rectangular Approximation:

$$\int f(x)\mathrm{d}x \approx \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}) \quad \text{or} \quad \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1}).$$

Trapezoid Approximation:

$$\int f(x)\mathrm{d}x \approx \sum_{i=1}^{n} \frac{1}{2}(f(x_i) + f(x_{i-1}))(x_i - x_{i-1}).$$

Bivariate Case: Consider integration of a function $f(x, y)$.

Suppose that both x and y are scalars.

Let x_0, x_1, x_2, $\cdots$, x_n be n nodes, which are sorted by order of size not necessarily equal intervals between x_{i-1} and x_i for $i = 1, 2, \cdots, n$.

Let y_0, y_1, y_2, $\cdots$, y_m be m nodes.

Rectangular Approximation:

$$\int\int f(x, y)\mathrm{d}x\mathrm{d}y \approx \sum_{i=1}^{n}\sum_{j=1}^{m} f(x_i, y_j)(x_i - x_{i-1})(y_j - y_{j-1}).$$

Trapezoid Approximation:

$$\int\int f(x.y)\mathrm{d}x\mathrm{d}y$$
$$\approx \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{1}{4}(f(x_i, y_j) + f(x_i, y_{j-1}) + f(x_{i-1}, y_j) + f(x_{i-1}, y_{j-1}))(x_i - x_{i-1})(y_j - y_{j-1}).$$

Applying to Bayes Method (Rectangular Approximation):

$$\begin{aligned}
\mathrm{E}(\theta|y) &= \frac{\int \theta f_{y|\theta}(y|\theta) f_\theta(\theta)\mathrm{d}\theta}{\int f_{y|\theta}(y|\theta) f_\theta(\theta)\mathrm{d}\theta} = \frac{\sum_{i=1}^n \theta_i f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)(\theta_i - \theta_{i-1})}{\sum_{i=1}^n f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)(\theta_i - \theta_{i-1})} \\
&= \frac{\sum_{i=1}^n \theta_i f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)}{\sum_{i=1}^n f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)} = \sum_{i=1}^n \theta_i \omega_i, \quad \text{for constant } \theta_i - \theta_{i-1},
\end{aligned}$$

where

$$\omega_i = \frac{f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)}{\sum_{i=1}^n f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)}.$$

Problem of Numerical Integration:

1. Choice of initial and terminal values $\Longrightarrow$ Truncation errors

2. Accumulation of computational errors by computer

3. Increase of computational burden for large dimension.

   $\Longrightarrow$ k dimension, and n nodes for each dimension $\Longrightarrow$ n^k

2.5.2 Evaluation of Expectation: Monte Carlo Integration

Univariate Case: Consider integration of a function $f(x)$.

Suppose that x is a scalar.

Let $x_1, x_2, \cdots, x_n$ be n random draws generated from $g(x)$.

$$\int f(x)\mathrm{d}x = \int \frac{f(x)}{g(x)} g(x)\mathrm{d}x = \mathrm{E}\Big(\frac{f(x)}{g(x)}\Big) \approx \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)}{g(x_i)}.$$

$\Longrightarrow$ **Importance Sampling (重点的サンプリング)**

Multivariate Case: Consider integration of a function $f(x)$.

Suppose that x is a vector.

Let $x_1, x_2, \cdots, x_n$ be n random draws generated from $g(x)$.

$$\int f(x)\mathrm{d}x = \int \frac{f(x)}{g(x)} g(x)\mathrm{d}x = \mathrm{E}\Big(\frac{f(x)}{g(x)}\Big) \approx \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)}{g(x_i)},$$

which is exacly the same as the univariate case.

Computational burden: $\Longrightarrow$ Univariate case: n, Multivariate case: n

Precision of integration ???

Especially, when $g(x)$ is not close to $f(x)$, approximation is prror.

Applying to Bayes Method:

$$\mathrm{E}(\theta|y) = \frac{\int \theta f_{y|\theta}(y|\theta) f_\theta(\theta)\mathrm{d}\theta}{\int f_{y|\theta}(y|\theta) f_\theta(\theta)\mathrm{d}\theta} = \frac{\displaystyle\int \theta \frac{f_{y|\theta}(y|\theta) f_\theta(\theta)}{g(\theta)} g(\theta)\mathrm{d}\theta}{\displaystyle\int \frac{f_{y|\theta}(y|\theta) f_\theta(\theta)}{g(\theta)} g(\theta)\mathrm{d}\theta} = \frac{(1/n)\sum_{i=1}^n \theta_i \omega(\theta_i)}{(1/n)\sum_{i=1}^n \omega(\theta_i)},$$

where

$$\omega(\theta_i) = \frac{f_{y|\theta}(y|\theta_i) f_\theta(\theta_i)}{g(\theta_i)}.$$

Choice of $g(\theta)$ — One Solution: Define $l(\theta) \equiv f_{y|\theta}(y|\theta) f_\theta(\theta)$.

$$\begin{aligned}\log l(\theta) &\approx \log l(\tilde\theta) + \frac{1}{l(\tilde\theta)} \frac{\partial l(\tilde\theta)}{\partial \theta}(\theta - \tilde\theta) \\ &\qquad + \frac{1}{2}(\theta - \tilde\theta)'\Big(-\frac{1}{l(\tilde\theta)^2} \frac{\partial l(\tilde\theta)}{\partial \theta} \frac{\partial l(\tilde\theta)}{\partial \theta'} + \frac{1}{l(\tilde\theta)} \frac{\partial^2 l(\tilde\theta)}{\partial\theta\partial\theta'}\Big)(\theta - \tilde\theta) \\ &= -\frac{1}{2}(\theta - \tilde\theta)'\Big(-\frac{1}{l(\tilde\theta)} \frac{\partial^2 l(\tilde\theta)}{\partial\theta\partial\theta'}\Big)(\theta - \tilde\theta), \qquad \text{when } \tilde\theta \text{ is a mode of } l(\theta).\end{aligned}$$

Thus, $N\Big(\tilde{\theta},\ \Big(-\frac{1}{l(\tilde{\theta})}\frac{\partial^2 l(\tilde{\theta})}{\partial\theta\partial\theta'}\Big)^{-1}\Big)$ might be taken as the importance density $g(\theta)$.

2.5.3 Evaluation of Expectation: Random Number Generation

Generate random draws of θ from the posterior distribution $f_{\theta|y}(\theta|y)$.

Then, $(1/n)\sum_{i=1}^{n}\theta_i$ is taken as a consistent estimator of $\mathrm{E}(\theta|y)$, where θ_i indicates the ith random draw generated from $f_{\theta|y}(\theta|y)$.

Note that $(1/n)\sum_{i=1}^{n}\theta_i \longrightarrow \mathrm{E}(\theta|y)$ under the condition $(1/n)\sum_{i=1}^{n}\theta_i < \infty$.

Bayesian confidence interval, median, quntiles and so on are obtained by sorting $\theta_1, \theta_2, \cdots, \theta_n$ in order of size.

$\Longrightarrow$ Sampling methods

2.6 Sampling Method I: Random Number Generation

Note that a lot of distribution functions are introduced in Kotz, Balakrishman and Johnson (2000a, 2000b, 2000c, 2000d, 2000e).

The random draws discussed in this section are based on uniform random draws between zero and one.

2.6.1 Uniform Distribution: $U(0, 1)$

Properties of Uniform Distribution: The most heuristic and simplest distribution is uniform.

The **uniform distribution** between zero and one is given by:

$$f(x) = \begin{cases} 1, & \text{for } 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Mean, variance and the moment-generating function are given by:

$$\mathrm{E}(X) = \frac{1}{2}, \qquad \mathrm{V}(X) = \frac{1}{12}, \qquad \phi(\theta) = \frac{e^{\theta} - 1}{\theta}.$$

Use L'Hospital's theorem to derive $\mathrm{E}(X)$ and $\mathrm{V}(X)$ using $\phi(\theta)$.

In the next section, we introduce an idea of generating uniform random draws, which in turn yield the other random draws by the transformation of variables, the inverse transform algorithm and so on.

Uniform Random Number Generators: It is no exaggeration to say that all the random draws are based on a uniform random number.

Once uniform random draws are generated, the various random draws such as exponential, normal, logistic, Bernoulli and other distributions are obtained by transforming the uniform random draws.

Thus, it is important to consider how to generate a uniform random number.

However, generally there is no way to generate exact uniform random draws.

As shown in Ripley (1987) and Ross (1997), a deterministic sequence that appears at random is taken as a sequence of random numbers.

First, consider the following relation:

$$m = k - [k/n]n,$$

where k, m and n are integers.

$[k/n]$ denotes the largest integer less than or equal to the argument.

In Fortran 77, it is written as `m=k-int(k/n)*n`, where $0 \leq m < n$.

m indicates the **remainder (余り)** when k is divided by n.

n is called the **modulus (商)**.

We define the right hand side in the equation above as:

$$k - [k/n]n \equiv k \bmod n.$$

Then, using the modular arithmetic we can rewrite the above equation as follows:

$$m = k \bmod n,$$

which is represented by: `m=mod(k,n)` in Fortran 77 and `m=k%n` in C language.

A basic idea of the uniform random draw is as follows.

Given x_{i-1}, x_i is generated by:

$$x_i = (ax_{i-1} + c) \bmod n,$$

where $0 \leq x_i < n$.

a and c are positive integers, called the **multiplier** and the **increment**, respectively.

The generator above have to be started by an initial value, which is called the **seed**.

$u_i = x_i/n$ is regarded as a uniform random number between zero and one.

This generator is called the **linear congruential generator (線形合同法)**.

Especially, when $c = 0$, the generator is called the **multiplicative linear congruential generator**.

This method was proposed by Lehmer in 1948 (see Lehmer, 1951).

If n, a and c are properly chosen, the period of the generator is n.

However, when they are not chosen very carefully, there may be a lot of serial correlation among the generated values.

Therefore, the performance of the congruential generators depend heavily on the choice of (a, c).

There is a great amount of literature on uniform random number generation.

See, for example, Fishman (1996), Gentle (1998), Kennedy and Gentle (1980), Law and Kelton (2000), Niederreiter (1992), Ripley (1987), Robert and Casella (1999), Rubinstein and Melamed (1998), Thompson (2000) and so on for the other congruential generators.

However, we introduce only two uniform random number generators.

Wichmann and Hill (1982 and corrigendum, 1984) describe a combination of three congruential generators for 16-bit computers.

The generator is given by:

$$x_i = 171 x_{i-1} \bmod 30269,$$
$$y_i = 172 y_{i-1} \bmod 30307,$$
$$z_i = 170 z_{i-1} \bmod 30323,$$

and

$$u_i = \left(\frac{x_i}{30269} + \frac{y_i}{30307} + \frac{z_i}{30323}\right) \bmod 1.$$

We need to set three seeds, i.e., x_0, y_0 and z_0, for this random number generator.

u_i is regarded as a uniform random draw within the interval between zero and one.

The period is of the order of 10^{12} (more precisely the period is 6.95×10^{12}).

For 32-bit computers, L'Ecuyer (1988) proposed a combination of k congruential generators that have prime moduli n_j, such that all values of $(n_j - 1)/2$ are relatively prime, and with multipliers that yield full periods.

Let the sequence from jth generator be $x_{j,1}, x_{j,2}, x_{j,3}, \cdots$.

Consider the case where each individual generator j is a maximum-period multiplicative linear congruential generator with modulus n_j and multiplier a_j, i.e.,

$$x_{j,i} \equiv a_j x_{j,i-1} \bmod n_j.$$

Assuming that the first generator is a relatively good one and that n_1 is fairly large, we form the ith integer in the sequence as:

$$x_i = \sum_{j=1}^{k} (-1)^{j-1} x_{j,i} \bmod (n_1 - 1),$$

where the other moduli n_j, $j = 2, 3, \cdots, k$, do not need to be large.

The normalization takes care of the possibility of zero occurring in this sequence:

$$u_i = \begin{cases} \dfrac{x_i}{n_1}, & \text{if } x_i > 0, \\ \dfrac{n_1 - 1}{n_1}, & \text{if } x_i = 0. \end{cases}$$

As for each individual generator j, note as follows.

Define $q = [n/a]$ and $r \equiv n \bmod a$, i.e., n is decomposed as $n = aq + r$, where $r < a$.

Therefore, for $0 < x < n$, we have:

$$\begin{aligned} ax \bmod n &= (ax - [x/q]n) \bmod n \\ &= \big(ax - [x/q](aq + r)\big) \bmod n \\ &= \big(a(x - [x/q]q) - [x/q]r\big) \bmod n \\ &= \big(a(x \bmod q) - [x/q]r\big) \bmod n. \end{aligned}$$

Practically, L'Ecuyer (1988) suggested combining two multiplicative congruential generators, where $k = 2$, $(a_1, n_1, q_1, r_1) = (40014, 2147483563, 53668, 12211)$ and (a_2, n_2, q_2, r_2) $= (40692, 2147483399, 52774, 3791)$ are chosen.

Two seeds are required to implement the generator.

The period of the generator proposed by L'Ecuyer (1988) is of the order of 10^{18} (more precisely 2.31×10^{18}), which is quite long and practically long enough.

L'Ecuyer (1988) presents the results of both theoretical and empirical tests, where the above generator performs well.

Furthermore, L'Ecuyer (1988) gives an additional portable generator for 16-bit computers. Also, see L'Ecuyer(1990, 1998).

To improve the length of period, the above generator proposed by L'Ecuyer (1988) is combined with the shuffling method suggested by Bays and Durham (1976), and it is introduced as `ran2` in Press, Teukolsky, Vetterling and Flannery (1992a, 1992b).

2.6.2 Transforming $U(0,1)$: Continuous Type

In this section, we focus on a continuous type of distributions, in which density functions are derived from the uniform distribution $U(0,1)$ by transformation of variables.

Normal Distribution: $N(0,1)$: The normal distribution with mean zero and variance one, i.e, the standard normal distribution, is represented by:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

for $-\infty < x < \infty$.

Mean, variance and the moment-generating function are given by:

$$\mathrm{E}(X) = 0, \qquad \mathrm{V}(X) = 1, \qquad \phi(\theta) = \exp\Big(\frac{1}{2}\theta^2\Big).$$

The normal random variable is constructed using two independent uniform random variables.

This transformation is well known as the Box-Muller (1958) transformation and is shown as follows.

Let U_1 and U_2 be uniform random variables between zero and one.

Suppose that U_1 is independent of U_2.

Consider the following transformation:

$$X_1 = \sqrt{-2\log(U_1)}\cos(2\pi U_2),$$
$$X_2 = \sqrt{-2\log(U_1)}\sin(2\pi U_2).$$

where we have $-\infty < X_1 < \infty$ and $-\infty < X_2 < \infty$ when $0 < U_1 < 1$ and $0 < U_2 < 1$.

Then, the inverse transformation is given by:

$$u_1 = \exp\left(-\frac{x_1^2 + x_2^2}{2}\right), \qquad u_2 = \frac{1}{2\pi}\arctan\frac{x_2}{x_1}.$$

We perform transformation of variables in multivariate cases.

From this transformation, the Jacobian is obtained as:

$$J = \begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} -x_1 \exp\Big(-\frac{1}{2}(x_1^2 + x_2^2)\Big) & -x_2 \exp\Big(-\frac{1}{2}(x_1^2 + x_2^2)\Big) \\ \frac{1}{2\pi} \frac{-x_2}{x_1^2 + x_2^2} & \frac{1}{2\pi} \frac{x_1}{x_1^2 + x_2^2} \end{vmatrix}$$

$$= -\frac{1}{2\pi} \exp\Big(-\frac{1}{2}(x_1^2 + x_2^2)\Big).$$

Let $f_x(x_1, x_2)$ be the joint density of X_1 and X_2 and $f_u(u_1, u_2)$ be the joint density of U_1 and U_2.

Since U_1 and U_2 are assumed to be independent, we have the following:

$$f_u(u_1, u_2) = f_1(u_1) f_2(u_2) = 1,$$

where $f_1(u_1)$ and $f_2(u_2)$ are the density functions of U_1 and U_2, respectively.

Note that $f_1(u_1) = f_2(u_2) = 1$ because U_1 and U_2 are uniform random variables between zero and one.

Accordingly, the joint density of X_1 and X_2 is:

$$
\begin{aligned}
f_x(x_1, x_2) &= |J| f_u\Big(\exp(-\frac{x_1^2 + x_2^2}{2}), \frac{1}{2\pi} \arctan \frac{x_2}{x_1}\Big) \\
&= \frac{1}{2\pi} \exp\Big(-\frac{1}{2}(x_1^2 + x_2^2)\Big) \\
&= \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2} x_1^2\Big) \times \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2} x_2^2\Big),
\end{aligned}
$$

which is a product of two standard normal distributions.

Thus, X_1 and X_2 are mutually independently distributed as normal random variables with mean zero and variance one.

See Hogg and Craig (1995, pp.177 – 178).

Therefore, to avoid computation of the sine, various algorithms have been invented (Ahrens and Dieter (1988), Fishman (1996), Gentle (1998), Marsaglia, MacLaren and Bray (1964) and so on).

Exponential Distribution: The exponential distribution with parameter β is written as:

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-\frac{x}{\beta}}, & \text{for } 0 < x < \infty, \\ 0, & \text{otherwise,} \end{cases}$$

for $\beta > 0$.

β indicates a scale parameter.

Mean, variance and the moment-generating function are obtained as follows:

$$\mathrm{E}(X) = \beta, \qquad \mathrm{V}(X) = \beta^2, \qquad \phi(\theta) = \frac{1}{1 - \beta\theta}.$$

The relation between the exponential random variable the uniform random variable is shown as follows:

When $U \sim U(0, 1)$, consider the following transformation:

$$X = -\beta \log(U).$$

Then, X is an exponential distribution with parameter β.

Because the transformation is given by $u = \exp(-x/\beta)$, the Jacobian is:

$$J = \frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{1}{\beta}\exp\Big(-\frac{1}{\beta}x\Big).$$

By transforming the variables, the density function of X is represented as:

$$f(x) = |J| f_u\Big(\exp(-\frac{1}{\beta}x)\Big) = \frac{1}{\beta}\exp\Big(-\frac{1}{\beta}x\Big),$$

where $f(\cdot)$ and $f_u(\cdot)$ denote the probability density functions of X and U, respectively.

Note that $0 < x < \infty$ because of $x = -\beta\log(u)$ and $0 < u < 1$.

Thus, the exponential distribution with parameter β is obtained from the uniform random draw between zero and one.

When $\beta = 2$, the exponential distribution reduces to the chi-square distribution with 2 degrees of freedom.

Gamma Distribution: $G(\alpha, \beta)$: The gamma distribution with parameters α and β, denoted by $G(\alpha, \beta)$, is represented as follows:

$$f(x) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & \text{for } 0 < x < \infty, \\ 0, & \text{otherwise,} \end{cases}$$

for $\alpha > 0$ and $\beta > 0$, where α is called a **shape parameter** and β denotes a scale parameter. $\Gamma(\cdot)$ is called the **gamma function**, which is the following function of α:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \,\mathrm{d}x.$$

The gamma function has the following features:

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha), \qquad \Gamma(1) = 1, \qquad \Gamma\Big(\frac{1}{2}\Big) = 2\Gamma\Big(\frac{3}{2}\Big) = \sqrt{\pi}.$$

Mean, variance and the moment-generating function are given by:

$$\mathrm{E}(X) = \alpha\beta, \qquad \mathrm{V}(X) = \alpha\beta^2, \qquad \phi(\theta) = \frac{1}{(1-\beta\theta)^\alpha}.$$

The gamma distribution with $\alpha = 1$ is equivalent to the exponential distribution shown in Section 2.6.2.

This fact is easily checked by comparing both moment-generating functions.

Now, utilizing the uniform random variable, the gamma distribution with parameters α and β are derived as follows.

The derivation shown in this section deals with the case where α is a positive integer, i.e., $\alpha = 1, 2, 3, \cdots$.

The random variables $Z_1, Z_2, \cdots, Z_\alpha$ are assumed to be mutually independently distributed as exponential random variables with parameter β, which are shown in Section 2.6.2.

Define $X = \sum_{i=1}^{\alpha} Z_i$.

Then, X has distributed as a gamma distribution with parameters α and β, where α should be an integer, which is proved as follows:

$$\phi_x(\theta) = \mathrm{E}(e^{\theta X}) = \mathrm{E}(e^{\theta \sum_{i=1}^{\alpha} Z_i}) = \prod_{i=1}^{\alpha} \mathrm{E}(e^{\theta Z_i}) = \prod_{i=1}^{\alpha} \phi_i(\theta) = \prod_{i=1}^{\alpha} \frac{1}{1-\beta\theta}$$
$$= \frac{1}{(1-\beta\theta)^{\alpha}},$$

where $\phi_x(\theta)$ and $\phi_i(\theta)$ represent the moment-generating functions of X and Z_i, respectively. Thus, sum of the α exponential random variables yields the gamma random variable with parameters α and β.

When α is large, we have serious problems computationally in the above algorithm, because α exponential random draws have to be generated to obtain one gamma random draw with parameters α and β.

When $\alpha = k/2$ and $\beta = 2$, the gamma distribution reduces to the chi-square distribution with k degrees of freedom.

Chi-Square Distribution: $\chi^2(k)$**:** The chi-square distribution with k degrees of freedom, denoted by $\chi^2(k)$, is written as follows:

$$f(x) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{1}{2}x}, & \text{for } 0 < x < \infty, \\ 0, & \text{otherwise,} \end{cases}$$

where k is a positive integer.

The chi-square distribution is equivalent to the gamma distribution with $\beta = 2$ and $\alpha = k/2$.

The chi-square distribution with $k = 2$ reduces to the exponential distribution with $\beta = 2$, shown in Section 2.6.2.

Mean, variance and the moment-generating function are given by:

$$\mathrm{E}(X) = k, \qquad \mathrm{V}(X) = 2k, \qquad \phi(\theta) = \frac{1}{(1-2\theta)^{k/2}}.$$

F Distribution: $F(m,n)$: The F distribution with m and n degrees of freedom, denoted by $F(m,n)$, is represented as:

$$f(x) = \begin{cases} \dfrac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\Big(\dfrac{m}{n}\Big)^{\frac{m}{2}} x^{\frac{m}{2}-1}\Big(1+\dfrac{m}{n}x\Big)^{-\frac{m+n}{2}}, & \text{for } 0 < x < \infty, \\ 0, & \text{otherwise,} \end{cases}$$

where m and n are positive integers.

Mean and variance are given by:

$$\mathrm{E}(X) = \frac{n}{n-2}, \qquad \text{for } n > 2,$$

$$\mathrm{V}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \qquad \text{for } n > 4.$$

The moment-generating function of F distribution does not exist.

One F random variable is derived from two chi-square random variables.

Suppose that U and V are independently distributed as chi-square random variables, i.e., $U \sim \chi^2(m)$ and $V \sim \chi^2(n)$.

Then, it is shown that $X = \dfrac{U/m}{V/n}$ has a F distribution with (m, n) degrees of freedom.

t **Distribution:** $t(k)$**:** The t distribution (or Student's t distribution) with k degrees of freedom, denoted by $t(k)$, is given by:

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{1}{\sqrt{k\pi}} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}},$$

for $-\infty < x < \infty$, where k does not have to be an integer but conventionally it is a positive integer.

When k is small, the t distribution has fat tails.

The t distribution with $k = 1$ is equivalent to the Cauchy distribution.

As k goes to infinity, the t distribution approaches the standard normal distribution, i.e., $t(\infty) = N(0, 1)$, which is easily shown by using the definition of e, i.e.,

$$\Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} = \Big(1 + \frac{1}{h}\Big)^{-\frac{hx^2+1}{2}} = \Big((1 + \frac{1}{h})^h\Big)^{-\frac{1}{2}x^2}\Big(1 + \frac{1}{h}\Big)^{-\frac{1}{2}} \longrightarrow e^{-\frac{1}{2}x^2},$$

where $h = k/x^2$ is set and h goes to infinity (equivalently, k goes to infinity).

Thus, a kernel of the t distribution is equivalent to that of the standard normal distribution.

Therefore, it is shown that as k is large the t distribution approaches the standard normal distribution.

Mean and variance of the t distribution with k degrees of freedom are obtained as:

$$\mathrm{E}(X) = 0, \qquad \text{for } k > 1,$$
$$\mathrm{V}(X) = \frac{k}{k-2}, \qquad \text{for } k > 2.$$

In the case of the t distribution, the moment-generating function does not exist, because all the moments do not necessarily exist.

For the t random variable X, we have the fact that $\mathrm{E}(X^p)$ exists when p is less than k.

Therefore, all the moments exist only when k is infinity.

One t random variable is obtained from chi-square and standard normal random variables.

Suppose that $Z \sim N(0, 1)$ is independent of $U \sim \chi^2(k)$.

Then, $X = Z/\sqrt{U/k}$ has a t distribution with k degrees of freedom.

Marsaglia (1984) gives a very fast algorithm for generating t random draws, which is based on a transformed acceptance/rejection method, which will be discussed later.

Gentle, J.E., 1998, *Random Number Generation and Monte Carlo Methods*, Springer-Verlag.

Kennedy, Jr. W.J. and Gentle, J.E., 1980, *Statistical Computing* (Statistics: Textbooks and Monographs, Vol.33), Marcel Dekker.

Ripley, B.D., 1987, *Stochastic Simulation*, John Wiley & Sons.