2.6 Functions of random variables

If ξ is a random variable, $y = g(x)$ a real function, then $\eta = g(\xi)$ is a function of ξ. Problems:

1. Is $\eta = g(\xi)$ a random variable?
2. If so, is there any connection between the distribution functions of ξ and η?

Notice for $\eta = g(\xi)$,

$$
\begin{aligned}
& \{\omega : \eta(\omega) \in B\} \\
= \; & \{\omega : g(\xi(\omega)) \in B\} \\
= \; & \Big\{\omega : \xi(\omega) \in \{x : g(x) \in B\}\Big\} \\
& B \in \mathcal{B}.
\end{aligned}
$$

Notice for $\eta = g(\xi)$,

$$\begin{aligned} & \{\omega : \eta(\omega) \in B\} \\ = & \{\omega : g(\xi(\omega)) \in B\} \\ = & \Big\{\omega : \xi(\omega) \in \{x : g(x) \in B\}\Big\} \\ & B \in \mathcal{B}. \end{aligned}$$

To require η being a random variable, it requires that $\Big\{\omega : \xi(\omega) \in \{x : g(x) \in B\}\Big\}$ is an event for any Borel set B. So, it is sufficient to require that for any Borel set B, $\{x : g(x) \in B\}$ is also a Borel set.

Definition

Suppose that $g(x)$ is a one dimensional real function, $\mathcal{B}$ is a Borel σ-field in $\mathbf{R}$. If for any $B \in \mathcal{B}$,

$$\{x : g(x) \in B\} \widehat{=} g^{-1}(B) \in \mathcal{B},$$

(that is, the pre-image under g of an arbitrary Borel set is also a Borel set) then we call $g(x)$ a Borel function.

Definition

Suppose that $g(x)$ is a one dimensional real function, $\mathcal{B}$ is a Borel σ-field in $\mathbf{R}$. If for any $B \in \mathcal{B}$,

$$\{x : g(x) \in B\} \widehat{=} g^{-1}(B) \in \mathcal{B},$$

(that is, the pre-image under g of an arbitrary Borel set is also a Borel set) then we call $g(x)$ a Borel function.

All piecewise continuous functions, piecewise monotone functions are Borel functions.

If ξ is a r.v. defined on the probability space $(\Omega, \mathcal{F}, P)$, $f(x)$ a Borel function. Let $\eta = f(\xi)$, then for an arbitrary $B \in \mathcal{B}$, we have

$$\begin{aligned} \{\omega : \eta(\omega) \in B\} &= \{\omega : f(\xi(\omega)) \in B\} \\ &= \{\omega : \xi(\omega) \in f^{-1}(B)\} \in \mathcal{F}, \end{aligned}$$

so η is a r.v.

If ξ is a r.v. defined on the probability space $(\Omega, \mathcal{F}, P)$, $f(x)$ a Borel function. Let $\eta = f(\xi)$, then for an arbitrary $B \in \mathcal{B}$, we have

$$\begin{aligned} \{\omega : \eta(\omega) \in B\} &= \{\omega : f(\xi(\omega)) \in B\} \\ &= \{\omega : \xi(\omega) \in f^{-1}(B)\} \in \mathcal{F}, \end{aligned}$$

so η is a r.v.

Similarly, if $f(x_1, \cdots, x_n)$ is a Borel function, then $\eta = f(\xi_1, \cdots, \xi_n)$ is a random variable.

2.5.1 Functions of discrete random variables

Example

Suppose that ξ has distribution sequence

$$\begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

Let $\eta = 2\xi - 1$,$\zeta = \xi^2$, find the distribution sequences of η and ζ.

Solution.

The distribution of $\eta = 2\xi - 1$ as follows:

$$\Bigg(\begin{matrix} -3 & -1 & 1 & 3 \\ & & & \end{matrix}$$

The distribution of $\eta = 2\xi - 1$ as follows:

$$\begin{pmatrix} -3 & -1 & 1 & 3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

The distribution of $\eta = 2\xi - 1$ as follows:

$$\begin{pmatrix} -3 & -1 & 1 & 3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

The distribution sequence of $\zeta = \xi^2$:

$$\left(\begin{matrix} 0 & 1 & 4 \\ & & \end{matrix} \right.$$

The distribution of $\eta = 2\xi - 1$ as follows:

$$\begin{pmatrix} -3 & -1 & 1 & 3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

The distribution sequence of $\zeta = \xi^2$:

$$\begin{pmatrix} 0 & 1 & 4 \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}.$$

In general, assume that ξ is such that

$$P(\xi = x_i) = p(x_i), \qquad i = 1, 2, \cdots,$$

then the distribution of $\eta = f(\xi)$ is

$$P(\eta = y_j) = \sum_{f(x_i) = y_j} p(x_i), \quad j = 1, 2, \cdots.$$

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k, \eta = r - k)$$

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k, \eta = r - k)$$

$$= \sum_{k=0}^{r} P(\xi = k) P(\eta = r - k)$$

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k, \eta = r - k)$$

$$= \sum_{k=0}^{r} P(\xi = k) P(\eta = r - k) = \sum_{k=0}^{r} \binom{n_1}{k} p^k q^{n_1 - k} \binom{n_2}{r-k} p^{r-k} q^{n_2 - r + k}$$

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k, \eta = r-k)$$

$$= \sum_{k=0}^{r} P(\xi = k) P(\eta = r-k) = \sum_{k=0}^{r} \binom{n_1}{k} p^k q^{n_1-k} \binom{n_2}{r-k} p^{r-k} q^{n_2-r+k}$$

$$= p^r q^{n_1+n_2-r} \sum_{k=0}^{r} \binom{n_1}{k} \binom{n_2}{r-k}$$

Example

Assume that $\xi \sim B(n_1, p)$, $\eta \sim B(n_2, p)$, and that ξ, η are independent. Find the distribution of $\zeta = \xi + \eta$.

Solution.

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k, \eta = r - k)$$

$$= \sum_{k=0}^{r} P(\xi = k) P(\eta = r - k) = \sum_{k=0}^{r} \binom{n_1}{k} p^k q^{n_1 - k} \binom{n_2}{r-k} p^{r-k} q^{n_2 - r + k}$$

$$= p^r q^{n_1 + n_2 - r} \sum_{k=0}^{r} \binom{n_1}{k} \binom{n_2}{r-k} = \binom{n_1 + n_2}{r} p^r q^{n_1 + n_2 - r}.$$

$\xi + \eta \sim B(n_1 + n_2, p)$.......the additivity property

$$\xi + \eta \sim B(n_1 + n_2, p)$$ ......the additivity property

The formula

$$P(\zeta = r) = \sum_{k=0}^{r} P(\xi = k) P(\eta = r - k).$$

is called the discrete convolution(卷积) formula.

2.5.2 Functions of continuous random variables

$\xi \sim$ pdf $p(x)$. $G(y)$ is the cdf of $\eta = f(\xi)$. That is,

$$G(y) = P(\eta \le y) = P(f(\xi) \le y).$$

Note that $D = \{x : f(x) \le y\}$ is a 1-dimensional Borel set, so

2.5.2 Functions of continuous random variables

$\xi \sim$ pdf $p(x)$. $G(y)$ is the cdf of $\eta = f(\xi)$. That is,

$$G(y) = P(\eta \le y) = P(f(\xi) \le y).$$

Note that $D = \{x : f(x) \le y\}$ is a 1-dimensional Borel set, so

$$G(y) = P(\xi \in D) = \int_{x \in D} p(x)dx.$$

Theorem

Suppose $f(x)$ is strictly monotone, and its inverse $f^{-1}(y)$ is continuously differentiable. Then $\eta = f(\xi)$ is a continuous random variable with density function:

$$g(y) = \begin{cases} p(f^{-1}(y))|(f^{-1}(y))'|, & y \in \text{ the range of } f(x), \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Without loss of generality, assume that $f(x)$ is strictly increasing, and $A < f(x) < B$ for $-\infty < x < \infty$.

Proof. Without loss of generality, assume that $f(x)$ is strictly increasing, and $A < f(x) < B$ for $-\infty < x < \infty$. Obviously, if $y \le A$ then $G(y) = 0$, so $g(y) = 0$. As $A < y < B$, $\{f(\xi) \le y\} = \{\xi \le f^{-1}(y)\}$, thus

Proof. Without loss of generality, assume that $f(x)$ is strictly increasing, and $A < f(x) < B$ for $-\infty < x < \infty$. Obviously, if $y \le A$ then $G(y) = 0$, so $g(y) = 0$. As $A < y < B$, $\{f(\xi) \le y\} = \{\xi \le f^{-1}(y)\}$, thus

$$G(y) = P(\eta \le y) = \int_{-\infty}^{f^{-1}(y)} p(x)dx.$$

Letting $x = f^{-1}(v)$, we have

Proof. Without loss of generality, assume that $f(x)$ is strictly increasing, and $A < f(x) < B$ for $-\infty < x < \infty$. Obviously, if $y \le A$ then $G(y) = 0$, so $g(y) = 0$. As $A < y < B$, $\{f(\xi) \le y\} = \{\xi \le f^{-1}(y)\}$, thus

$$G(y) = P(\eta \le y) = \int_{-\infty}^{f^{-1}(y)} p(x)dx.$$

Letting $x = f^{-1}(v)$, we have

$$G(y) = \int_{A}^{y} p(f^{-1}(v))(f^{-1}(v))'dv = \int_{-\infty}^{y} g(v)dv.$$

As $y \ge B$, $G(y) = 1$, so $g(y) = 0$. $\square$

Corollary

If $y = f(x)$ is piecewise strictly monotone in disjoint intervals $I_1, I_2, \cdots$, and its inverse $h_i(y)$ in the i-th interval is continuously differentiable. Then $\eta = f(\xi)$ is a continuous random variable, whose density is

$$g(y) = \begin{cases} \sum p(h_i(y))|h_i'(y)|, & \\ \qquad y \in \text{ the definition domain of each } h_i, & \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Let $E_i(y) = \{x : f(x) \le y, x \in I_i\}$. Observe that $\{f(\xi) \le y\} = \{\xi \in \sum_i E_i(y)\}$. We obtain

$$P(\eta \le y) = P(\xi \in \sum_i E_i(y)) = \sum_i \int_{E_i(y)} p(x)dx$$

Proof. Let $E_i(y)=\{x: f(x)\le y, x\in I_i\}$. Observe that $\{f(\xi)\le y\}=\{\xi\in\sum_i E_i(y)\}$. We obtain

$$\begin{aligned} P(\eta\le y) &= P(\xi\in\sum_i E_i(y))=\sum_i\int_{E_i(y)}p(x)dx \\ &= \sum_i\int_{-\infty}^{y}p(h_i(u))|h_i'(u)|du \end{aligned}$$

Proof. Let $E_i(y) = \{x : f(x) \le y, x \in I_i\}$. Observe that $\{f(\xi) \le y\} = \{\xi \in \sum_i E_i(y)\}$. We obtain

$$\begin{aligned} P(\eta \le y) &= P(\xi \in \sum_i E_i(y)) = \sum_i \int_{E_i(y)} p(x)dx \\ &= \sum_i \int_{-\infty}^{y} p(h_i(u))|h_i'(u)|du \\ &= \int_{-\infty}^{y} \sum_i p(h_i(x))|h_i'(x)|dx, \end{aligned}$$

Example

Assume $\xi \sim N(0,1)$, calculate the density function of $\eta = \xi^2$.

Example

Assume $\xi \sim N(0,1)$, calculate the density function of $\eta = \xi^2$.

Solution. Obviously, when $y \le 0$, $F_\eta(y) = P(\xi^2 \le y) = 0$.

Example

Assume $\xi \sim N(0,1)$, calculate the density function of $\eta = \xi^2$.

Solution. Obviously, when $y \le 0$, $F_\eta(y) = P(\xi^2 \le y) = 0$. When $y > 0$,

$$F_\eta(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \varphi(u)du,$$

Example

Assume $\xi \sim N(0,1)$, calculate the density function of $\eta = \xi^2$.

Solution. Obviously, when $y \leq 0$, $F_\eta(y) = P(\xi^2 \leq y) = 0$. When $y > 0$,

$$F_\eta(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \varphi(u)du,$$

and then

$$\begin{aligned} p_\eta(y) =& \varphi(\sqrt{y})(\sqrt{y})' - \varphi(-\sqrt{y})(-\sqrt{y})' \\ =& \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}. \end{aligned}$$

Example

Assume $\theta \sim U[0,1]$ and a function $F(x)$ possesses the same three properties required of a distribution function. Calculate the distribution of $\xi = F^{-1}(\theta)$, where $F^{-1}(y) = \sup\{x : F(x) < y\}$.

我们称$F^{-1}(y)=\sup\{x: F(x)<y\}$为分布函数$F(x)$的广义反函数，根据上确界的定义和分布函数的性质可以验证广义反函数有如下性质：

(i) $F^{-1}(y)$ $(0<y<1)$是y的单调不减函数；

(ii) $F(F^{-1}(y))\geq y$. 若$F(x)$在$x=F^{-1}(y)$处连续，则$F(F^{-1}(y))=y$;

(iii) $F^{-1}(y)\leq x$的充分必要条件是$y\leq F(x)$.

Solution. By the properties of F^{-1}, we have $F^{-1}(y) \le x \Leftrightarrow y \le F(x)$.

$$P(\xi \le x) = P(F^{-1}(\theta) \le x)$$

Solution. By the properties of F^{-1}, we have $F^{-1}(y) \le x \Leftrightarrow y \le F(x)$.

$$\begin{aligned} P(\xi \le x) &= P(F^{-1}(\theta) \le x) \\ &= P(\theta \le F(x)) \end{aligned}$$

Solution. By the properties of F^{-1}, we have $F^{-1}(y) \le x \Leftrightarrow y \le F(x)$.

$$\begin{aligned} P(\xi \le x) &= P(F^{-1}(\theta) \le x) \\ &= P(\theta \le F(x)) \\ &= F(x). \end{aligned}$$

Example

Assume that ξ has a continuous cdf $F(x)$, calculate the cdf of $\theta = F(\xi)$.

Solution.

Example

Assume that ξ has a continuous cdf $F(x)$, calculate the cdf of $\theta = F(\xi)$.

Solution. Obviously, when $y < 0$, $P(\theta \le y) = 0$; when $y \ge 1$, $P(\theta \le y) = 1$.

Example

Assume that ξ has a continuous cdf $F(x)$, calculate the cdf of $\theta = F(\xi)$.

Solution. Obviously, when $y < 0$, $P(\theta \le y) = 0$; when $y \ge 1$, $P(\theta \le y) = 1$. Now assume $0 \le y < 1$. Define for $0 \le y \le 1$

$$F^{-1}(y) = \sup\{x : F(x) < y\}.$$

This is the inverse of $F(x)$.

Example

Assume that ξ has a continuous cdf $F(x)$, calculate the cdf of $\theta = F(\xi)$.

Solution. Obviously, when $y < 0$, $P(\theta \le y) = 0$; when $y \ge 1$, $P(\theta \le y) = 1$. Now assume $0 \le y < 1$. Define for $0 \le y \le 1$

$$F^{-1}(y) = \sup\{x : F(x) < y\}.$$

This is the inverse of $F(x)$. Thus we have

$$\begin{aligned} P(\theta \le y) &= P(F(\xi) \le y) = P(\xi \le F^{-1}(y)) \\ &= F(F^{-1}(y)) = y. \end{aligned}$$

Example

Assume that ξ has a continuous cdf $F(x)$, calculate the cdf of $\theta = F(\xi)$.

Solution. Obviously, when $y < 0$, $P(\theta \leq y) = 0$; when $y \geq 1$, $P(\theta \leq y) = 1$. Now assume $0 \leq y < 1$. Define for $0 \leq y \leq 1$

$$F^{-1}(y) = \sup\{x : F(x) < y\}.$$

This is the inverse of $F(x)$. Thus we have

$$\begin{aligned} P(\theta \leq y) &= P(F(\xi) \leq y) = P(\xi \leq F^{-1}(y)) \\ &= F(F^{-1}(y)) = y. \end{aligned}$$

2.5.3 Functions of continuous random vectors

$(\xi_1, \cdots, \xi_n) \sim$ pdf $p(x_1, \cdots, x_n)$.

Let $\eta = f(\xi_1, \cdots, \xi_n)$, then the distribution function of η is determined by the following

$$F_\eta(y) = P(f(\xi_1, \cdots, \xi_n) \leq y)$$

2.5.3 Functions of continuous random vectors

$(\xi_1, \cdots, \xi_n) \sim$ pdf $p(x_1, \cdots, x_n)$.

Let $\eta = f(\xi_1, \cdots, \xi_n)$, then the distribution function of η is determined by the following

$$\begin{aligned} F_\eta(y) &= P(f(\xi_1, \cdots, \xi_n) \le y) \\ &= \int \cdots \int_{f(x_1, \cdots, x_n) \le y} p(x_1, \cdots, x_n) dx_1 \cdots dx_n. \end{aligned}$$

Special cases: **1.** $\eta = \xi_1 + \xi_2$

Special cases: **1.** $\eta = \xi_1 + \xi_2$

$$F_\eta(y) = \int\int_{x_1+x_2\le y} p(x_1, x_2)dx_1dx_2$$

Special cases: **1.** $\eta = \xi_1 + \xi_2$

$$\begin{aligned} F_\eta(y) &= \int\int_{x_1+x_2\le y} p(x_1,x_2)dx_1dx_2 \\ &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{y-x_1} p(x_1,x_2)dx_2 \end{aligned}$$

Special cases: **1.** $\eta = \xi_1 + \xi_2$

$$
\begin{aligned}
F_\eta(y) &= \iint_{x_1+x_2\le y} p(x_1,x_2)dx_1dx_2 \\
&= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{y-x_1} p(x_1,x_2)dx_2 \\
&\overset{x_2=z-x_1}{=} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{y} p(x_1,z-x_1)dz \\
&= \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} p(x_1,z-x_1)dx_1\right)dz.
\end{aligned}
$$

Special cases: 1. $\eta = \xi_1 + \xi_2$

$$
\begin{aligned}
F_\eta(y) &= \iint_{x_1+x_2\le y} p(x_1,x_2)dx_1dx_2 \\
&= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{y-x_1} p(x_1,x_2)dx_2 \\
&\overset{x_2=z-x_1}{=} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{y} p(x_1,z-x_1)dz \\
&= \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} p(x_1,z-x_1)dx_1\right)dz.
\end{aligned}
$$

$$p_\eta(y) = \int_{-\infty}^{\infty} p(x_1, y-x_1)dx_1.$$

When $\xi_1 \sim pdf\ p_1(x)$ and $\xi_2 \sim pdf\ p_2(x)$ are independent, the pdf of $\xi_1 + \xi_2$ is

$$p_\eta(z) = \int_{-\infty}^{\infty} p_1(x) p_2(z-x) dx.$$

When $\xi_1 \sim pdf\ p_1(x)$ and $\xi_2 \sim pdf\ p_2(x)$ are independent, the pdf of $\xi_1 + \xi_2$ is

$$p_\eta(z) = \int_{-\infty}^{\infty} p_1(x)p_2(z-x)dx.$$

Similarly, we have

$$p_\eta(z) = \int_{-\infty}^{\infty} p_1(z-x)p_2(x)dx.$$

When $\xi_1 \sim pdf\ p_1(x)$ and $\xi_2 \sim pdf\ p_2(x)$ are independent, the pdf of $\xi_1 + \xi_2$ is

$$p_\eta(z) = \int_{-\infty}^{\infty} p_1(x) p_2(z-x) dx.$$

Similarly, we have

$$p_\eta(z) = \int_{-\infty}^{\infty} p_1(z-x) p_2(x) dx.$$

Convolution formulas

Example

Suppose that ξ, η are i.i.d.r.v.s $\sim N(0,1)$. Calculate the pdf of $\zeta = \xi + \eta$.

Example

Suppose that ξ, η are i.i.d.r.v.s $\sim N(0,1)$. Calculate the pdf of $\zeta = \xi + \eta$.

Solution. For an arbitrary $z \in \mathbf{R}$,

$$p_\zeta(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} dx$$

Example

Suppose that ξ, η are i.i.d.r.v.s $\sim N(0,1)$. Calculate the pdf of $\zeta = \xi + \eta$.

Solution. For an arbitrary $z \in \mathbf{R}$,

$$
\begin{aligned}
p_\zeta(z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-(\sqrt{2}x - \frac{z}{\sqrt{2}})^2/2} dx
\end{aligned}
$$

Example

Suppose that ξ, η are i.i.d.r.v.s $\sim N(0,1)$. Calculate the pdf of $\zeta = \xi + \eta$.

Solution. For an arbitrary $z \in \mathbf{R}$,

$$
\begin{aligned}
p_\zeta(z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-(\sqrt{2}x - \frac{z}{\sqrt{2}})^2/2} dx = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{z^2}{4}},
\end{aligned}
$$

which implies $\zeta = \xi + \eta \sim N(0,2)$.

In general , if ξ, η are indept., and $\xi \sim N(a, \sigma_1^2)$, $\eta \sim N(b, \sigma_2^2)$, then $\xi + \eta \sim N(a + b, \sigma_1^2 + \sigma_2^2)$.

$$\xi_i \sim N(\mu_i, \sigma_i^2), i = 1, \cdots, n, \text{ indept.} \Longrightarrow$$
$$\xi_1 + \cdots + \xi_n \sim N(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2).$$

Proof. Let

$$c = \frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} = \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}.$$

We have

$$p_\xi(z-y)p_\eta(y) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(z-y-a)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-b)^2}{2\sigma_2^2}}$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(z-a)^2}{2\sigma_1^2} - \frac{y^2}{2\sigma_1^2} + 2y\frac{z-a}{2\sigma_1^2} - \frac{y^2}{2\sigma_2^2} - \frac{b^2}{2\sigma_2^2} + 2y\frac{b}{2\sigma_2^2}}$$

$$= e^{-\frac{(z-a-b)^2}{2(\sigma_1^2+\sigma_2^2)}} \frac{1}{2\pi\sigma_1\sigma_2} e^{-c\left(y - \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}(z-a) - \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}b\right)^2}.$$

It follows that

$$p_{\xi+\eta}(z) = \int_{-\infty}^{\infty} p_\xi(z-y)p_\eta(y)dy$$
$$= C_0 e^{-\frac{(z-a-b)^2}{2(\sigma_1^2+\sigma_2^2)}} = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{(z-a-b)^2}{2(\sigma_1^2+\sigma_2^2)}}.$$

So, $\xi + \eta \sim N(a+b, \sigma_1^2+\sigma_2^2)$.

Example

Suppose that ξ, η are indept. with the following density functions:

$$p_\xi(x) = \begin{cases} ae^{-ax}, & x > 0, \\ 0, & x \le 0, \end{cases} \qquad a > 0,$$

and

$$p_\eta(x) = \begin{cases} be^{-bx}, & x > 0, \\ 0, & x \le 0, \end{cases} \qquad b > 0.$$

Calculate the density function of $\zeta = \xi + \eta$.

Solution.

Solution. Observe that $p_\xi(x)p_\eta(z-x) \neq 0$ iff $x > 0$ and $z - x > 0$ iff $z > x > 0$.

Solution. Observe that $p_\xi(x)p_\eta(z-x) \neq 0$ iff $x > 0$ and $z - x > 0$ iff $z > x > 0$. Hence, $p_\zeta(z) = \int_{-\infty}^{\infty} p_\xi(x)p_\eta(z-x)dx = 0$ when $z \leq 0$;

Solution. Observe that $p_\xi(x)p_\eta(z-x) \neq 0$ iff $x > 0$ and $z - x > 0$ iff $z > x > 0$. Hence, $p_\zeta(z) = \int_{-\infty}^{\infty} p_\xi(x)p_\eta(z-x)dx = 0$ when $z \leq 0$;

$$p_\zeta(z) = \int_0^z ae^{-ax}be^{-b(z-x)}dx = abe^{-bz}\int_0^z e^{-(a-b)x}dx$$

when $z > 0$.

We take the following two cases into account:

Solution. Observe that $p_\xi(x)p_\eta(z-x)\neq 0$ iff $x>0$ and $z-x>0$ iff $z>x>0$. Hence, $p_\zeta(z)=\int_{-\infty}^{\infty}p_\xi(x)p_\eta(z-x)dx=0$ when $z\leq 0$;

$$p_\zeta(z)=\int_0^z ae^{-ax}be^{-b(z-x)}dx=abe^{-bz}\int_0^z e^{-(a-b)x}dx$$

when $z>0$.

We take the following two cases into account:

(1) If $a=b$, then $p_\zeta(z)=abze^{-bz}$;

Solution. Observe that $p_\xi(x)p_\eta(z-x) \neq 0$ iff $x > 0$ and $z - x > 0$ iff $z > x > 0$. Hence, $p_\zeta(z) = \int_{-\infty}^{\infty} p_\xi(x)p_\eta(z-x)dx = 0$ when $z \le 0$;

$$p_\zeta(z) = \int_0^z ae^{-ax}be^{-b(z-x)}dx = abe^{-bz}\int_0^z e^{-(a-b)x}dx$$

when $z > 0$.

We take the following two cases into account:

(1) If $a = b$, then $p_\zeta(z) = abze^{-bz}$;

(2) If $a \neq b$, then

$$p_\zeta(z) = \frac{ab}{a-b}(e^{-bz} - e^{-az}).$$

2. $\eta = \xi_1/\xi_2$

2. $\eta = \xi_1/\xi_2$

$$\begin{aligned} F_\eta(y) &= P(\frac{\xi_1}{\xi_2} \le y) = \int\int_{x_1/x_2 \le y} p(x_1,x_2)dx_1dx_2 \\ &= \int_0^\infty dx_2 \int_{-\infty}^{yx_2} p(x_1,x_2)dx_1 \\ &+ \int_{-\infty}^0 dx_2 \int_{yx_2}^\infty p(x_1,x_2)dx_1. \end{aligned}$$

Letting $x_1 = zx_2$ and noticing $z = -\infty$ when $x_1 = \infty$ and $x_2 < 0$, we obtain

$$\begin{aligned} F_\eta(y) &= \int_0^\infty dx_2 \int_{-\infty}^y p(zx_2, x_2)x_2 dz \\ &+ \int_{-\infty}^0 dx_2 \int_y^{-\infty} p(zx_2, x_2)x_2 dz \\ &= \int_0^\infty dx_2 \int_{-\infty}^y p(zx_2, x_2)x_2 dz \\ &- \int_{-\infty}^0 dx_2 \int_{-\infty}^y p(zx_2, x_2)x_2 dz. \end{aligned}$$

and exchanging the order of integration,

$$\begin{aligned} F_\eta(y) &= \int_{-\infty}^{y} [\int_0^{\infty} p(zx_2, x_2)x_2 dx_2 \\ &\quad - \int_{-\infty}^{0} p(zx_2, x_2)x_2 dx_2] dz \\ &= \int_{-\infty}^{y} p_\eta(z) dz. \end{aligned}$$

This shows that $\eta = \xi_1/\xi_2$ has the density function

$$p_\eta(z) = \int_{-\infty}^{\infty} p(zx, x)|x| dx.$$

Example

Suppose that ξ and η are independent standard normal random variables. Find the distribution of $\zeta = \xi/\eta$.

Example

Suppose that ξ and η are independent standard normal random variables. Find the distribution of $\zeta = \xi/\eta$.

Solution. We have

$$p_\zeta(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(zx)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} |x| dx$$

$$= \int_0^{\infty} \frac{1}{\pi} e^{-\frac{(z^2+1)x^2}{2}} x dx = \frac{1}{\pi(z^2+1)}.$$

Example

Suppose that ξ, η are independent identically distributed random variables with a common distribution $U(0, a)$. Calculate the density function of ξ/η.

Solution. Observe that

$$p_\xi(x) = p_\eta(x) = \begin{cases} \frac{1}{a}, & 0 \le x \le a, \\ 0, & \text{otherwise.} \end{cases}$$

Since ξ, η are indept, only when $0 \le xz \le a$ and $0 \le x \le a$

$$p(zx, x) = p_\xi(zx)p_\eta(x) = \frac{1}{a^2} \ne 0.$$

When $z < 0$, it follows that for any x

$$p(zx, x) = 0,$$

which implies that $p_{\xi/\eta}(z) = 0$;

when $0 \le z < 1$, it follows obviously $0 \le xz \le a$, so we have

$$p_{\xi/\eta}(z) = \int_0^a \frac{1}{a^2} x dx = \frac{1}{2}.$$

When $z \ge 1$, the integral becomes

$$p_{\xi/\eta}(z) = \int_0^{a/z} \frac{1}{a^2} x dx = \frac{1}{2z^2}.$$

3. Distributions of order statistics

$\xi_1, \cdots, \xi_n$ are independent identically distributed random variables with the common distribution function $F(x)$.

Order statistics:

$$\xi_1^* \le \cdots \le \xi_n^*.$$

3. Distributions of order statistics

$\xi_1, \cdots, \xi_n$ are independent identically distributed random variables with the common distribution function $F(x)$.

Order statistics:

$$\xi_1^* \leq \cdots \leq \xi_n^*.$$

$$\xi_1^* = \min\{\xi_1, \cdots, \xi_n\},\ \xi_n^* = \max\{\xi_1, \cdots, \xi_n\}.$$

(1) The distribution of ξ_n^*

(1) The distribution of ξ_n^*

$$P(\xi_n^* \le x)$$

(1) The distribution of ξ_n^*

$$P(\xi_n^* \le x) = P(\xi_1 \le x, \xi_2 \le x, \cdots, \xi_n \le x)$$

(1) The distribution of ξ_n^*

$$
\begin{aligned}
P(\xi_n^* \le x) &= P(\xi_1 \le x, \xi_2 \le x, \cdots, \xi_n \le x) \\
&= P(\xi_1 \le x) P(\xi_2 \le x) \cdots P(\xi_n \le x) \\
&= [F(x)]^n.
\end{aligned}
$$

(2) The distributions of ξ_1^*

(2) The distributions of ξ_1^*

For this, we consider the complement event $\{\xi_1^* > x\}$ of $\{\xi_1^* \le x\}$.

$$P(\xi_1^* > x)$$

(2) The distributions of ξ_1^*

For this, we consider the complement event $\{\xi_1^* > x\}$ of $\{\xi_1^* \le x\}$.

$$P(\xi_1^* > x) \;=\; P(\xi_1 > x, \xi_2 > x, \cdots, \xi_n > x)$$

(2) The distributions of ξ_1^*

For this, we consider the complement event $\{\xi_1^* > x\}$ of $\{\xi_1^* \le x\}$.

$$\begin{aligned} P(\xi_1^* > x) &= P(\xi_1 > x, \xi_2 > x, \cdots, \xi_n > x) \\ &= P(\xi_1 > x)P(\xi_2 > x)\cdots P(\xi_n > x) \\ &= [1 - F(x)]^n. \end{aligned}$$

(2) The distributions of ξ_1^*

For this, we consider the complement event $\{\xi_1^* > x\}$ of $\{\xi_1^* \le x\}$.

$$\begin{aligned} P(\xi_1^* > x) &= P(\xi_1 > x, \xi_2 > x, \cdots, \xi_n > x) \\ &= P(\xi_1 > x)P(\xi_2 > x)\cdots P(\xi_n > x) \\ &= [1 - F(x)]^n. \end{aligned}$$

Hence we have

$$P(\xi_1^* \le x) = 1 - [1 - F(x)]^n.$$

(3) The distribution of ξ_k^*

(3) The distribution of ξ_k^*

$$F_k(x) \; = \; P(\xi_k^* \le x) = P\left(\sharp\{i : \xi_i \le x\} \ge k\right)$$

(3) The distribution of ξ_k^*

$$\begin{aligned} F_k(x) &= P(\xi_k^* \le x) = P\left(\sharp\{i : \xi_i \le x\} \ge k\right) \\ &= \sum_{j=k}^{n} P\left(\sharp\{i : \xi_i \le x\} = j\right) \end{aligned}$$

(3) The distribution of ξ_k^*

$$\begin{aligned} F_k(x) &= P(\xi_k^* \le x) = P\left(\sharp\{i : \xi_i \le x\} \ge k\right) \\ &= \sum_{j=k}^{n} P\left(\sharp\{i : \xi_i \le x\} = j\right) \\ &= \sum_{j=k}^{n} \binom{n}{j} F^j(x)[1 - F(x)]^{n-j}. \end{aligned}$$

特别地，当$\xi_1,\dots,\xi_n \sim U[0,1]$时

$$F_k(x)=\sum_{j=k}^{n}\binom{n}{j}x^j(1-x)^{n-j},\ \ 0\le x\le 1.$$

密度为

$$p_k(x)=F_k'(x)=\frac{n!}{(k-1)!(n-k)!}x^{k-1}(1-x)^{n-k},\ \ 0\le x\le 1.$$

即

$$\xi_k^*\sim Beta(k,n-k+1).$$

一般地，当$\xi_1, \ldots, \xi_n \sim F(x)$时

$$F_k(x) = \frac{n!}{(k-1)!(n-k)!} \int_0^{F(x)} u^{k-1}(1-u)^{n-k} du.$$

密度为

$$p_k(x) = \frac{n!}{(k-1)!(n-k)!} F^{k-1}(x)[1-F(x)]^{n-k} p(x).$$

(4) The joint distribution of (ξ_1^*, ξ_n^*)

(4) The joint distribution of (ξ_1^*, ξ_n^*)

$$F_{1,n}(x, y) \;=\; P(\xi_1^* \le x, \xi_n^* \le y)$$

(4) The joint distribution of (ξ_1^*, ξ_n^*)

$$\begin{aligned} F_{1,n}(x,y) &= P(\xi_1^* \le x, \xi_n^* \le y) \\ &= P(\xi_n^* \le y) - P(\xi_1^* > x, \xi_n^* \le y) \\ &= [F(y)]^n - P\Big(\bigcap_{i=1}^n (x < \xi_i \le y)\Big). \end{aligned}$$

(4) The joint distribution of (ξ_1^*, ξ_n^*)

$$
\begin{aligned}
F_{1,n}(x,y) &= P(\xi_1^* \le x, \xi_n^* \le y) \\
&= P(\xi_n^* \le y) - P(\xi_1^* > x, \xi_n^* \le y) \\
&= [F(y)]^n - P(\bigcap_{i=1}^{n}(x < \xi_i \le y)).
\end{aligned}
$$

So, when $x < y$

$$F(x,y) = [F(y)]^n - [F(y) - F(x)]^n$$

and when $x \ge y$

$$F_{1,n}(x,y) = [F(y)]^n.$$

Suppose that $F(x)$ has density $p(x)$.

Suppose that $F(x)$ has density $p(x)$. The density of (ξ_1^*, ξ_n^*) is

$$p_{1,n}(x,y) = \frac{\partial^2 F_{1,n}(x,y)}{\partial y \partial x}$$

Suppose that $F(x)$ has density $p(x)$. The density of (ξ_1^*, ξ_n^*) is

$$
\begin{aligned}
p_{1,n}(x,y) &= \frac{\partial^2 F_{1,n}(x,y)}{\partial y \partial x} \\
&= \begin{cases} n(n-1)[F(y)-F(x)]^{n-2} p(x)p(y), & x < y, \\ 0, & x \geq y. \end{cases}
\end{aligned}
$$

概率微元法求密度函数: 不妨设$x < y$. 由于

$$P(x < \xi_1^* \le x + dx, y < \xi_n^* \le y + dy) = p_{1,n}(x, y)dxdy + o(dxdy).$$

右边是左边概率的主要部分,即概率微元.

对充分小的微元$dxdy$, 事件$\{x < \xi_1^* \le x + dx, y < \xi_n^* \le y + dy\}$ 意味着: $\xi_1, \cdots, \xi_n$ 中有

- 1个落在$(x, x + dx]$内, (每个观察值落在这个区间的概率为$F(x + dx) - F(x) \approx p(x)dx$);
- $n - 2$个观察值落在$(x + dx, y]$内, (每个观察值落在这个区间的概率为$F(y) - F(x + dx) \approx F(y) - F(x)$);
- 1个观察值落在$(y, y + dy]$内,(每个观察值落在这个区间的概率为$F(y + dy) - F(y) \approx p(y)dy$).

所以

$$
\begin{aligned}
&P(x < \xi_1^* \le x + dx, y < \xi_n^* \le y + dy) \\
=&\frac{n!}{1!(n-2)!1!} p(x)dx[F(y) - F(x)]^{n-2} p(y)dy + o(dxdy) \\
=&n(n-1)[F(y) - F(x)]^{n-2} p(x)p(y)dxdy + o(dxdy).
\end{aligned}
$$

所以

$$
\begin{aligned}
&P(x < \xi_1^* \le x + dx, y < \xi_n^* \le y + dy) \\
=&\frac{n!}{1!(n-2)!1!} p(x)dx[F(y) - F(x)]^{n-2} p(y)dy + o(dxdy) \\
=&n(n-1)[F(y) - F(x)]^{n-2} p(x)p(y)dxdy + o(dxdy).
\end{aligned}
$$

这样

$$
p_{1,n}(x,y)dxdy = n(n-1)[F(y) - F(x)]^{n-2} p(x)p(y)dxdy.
$$

从而

$$
p_{1,n}(x,y) = n(n-1)[F(y) - F(x)]^{n-2} p(x)p(y), \quad x \le y.
$$

用概率微元法可以求得(ξ_i^*, ξ_j^*) $(i < j)$的密度为

$$
\begin{aligned}
p_{i,j}(y_i, y_j) = & \frac{n!}{(i-1)!(j-i-1)!(n-j)!} p(y_i) p(y_j) \\
& \times F^{i-1}(y_i) (F(y_j) - F(y_i))^{j-i-1} (1 - F(y_j))^{n-j}, \\
& \quad y_i \le y_j.
\end{aligned}
$$

$(\xi_1^*, \ldots, \xi_n^*)$的密度为

$$g(y_1, \ldots, y_n) = n! p(y_1) \cdots p(y_n), y_1 \leq y_2 \leq \ldots \leq y_n.$$

(5) The joint distribution of $R = \xi_n^* - \xi_1^*$

(5) The joint distribution of $R = \xi_n^* - \xi_1^*$

$$p_R(r) = \int_{-\infty}^{\infty} p_{1,n}(x, x+r)dx$$

(5) The joint distribution of $R = \xi_n^* - \xi_1^*$

$$
\begin{aligned}
p_R(r) &= \int_{-\infty}^{\infty} p_{1,n}(x, x+r)dx \\
&= \int_{-\infty}^{\infty} n(n-1)[F(x+r) - F(x)]^{n-2} p(x)p(x+r)dx.
\end{aligned}
$$

2.5.4 Transforms of random vectors

$$(\xi_1, \cdots, \xi_n) \sim \text{ pdf } p(x_1, \cdots, x_n)$$

and

$$\begin{array}{ll} y_1 = f_1(x_1, \cdots, x_n), & \\ \cdots & \text{measurable functions.} \\ y_m = f_m(x_1, \cdots, x_n) & \end{array}$$

Let $\eta_1 = f_1(\xi_1, \cdots, \xi_n), \cdots, \eta_m = f_m(\xi_1, \cdots, \xi_n)$. Then $(\eta_1, \cdots, \eta_m)$ is a random vector and its cdf is

$$\begin{aligned} G(y_1, \cdots, y_m) &= P(\eta_1 \le y_1, \cdots, \eta_m \le y_m) \\ &= \int \cdots \int_D p(x_1, \cdots, x_n) dx_1 \cdots dx_n, \end{aligned}$$

where D is an n-dimensional domain:

$$\{(x_1, \cdots, x_n) : \quad f_1(x_1, \cdots, x_n) \le y_1, \quad \cdots, \quad f_m(x_1, \cdots, x_n) \le y_m\} \quad .$$

Theorem

If $m = n$, $f_j, j = 1, \cdots, n$ *have unique inverse functions* $x_i = x_i(y_1, \cdots, y_n), i = 1, \cdots, n$, *and*

$$J = \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)} \neq 0.$$

Then $(\eta_1, \cdots, \eta_n)$ *has density function* $q(y_1, \cdots, y_n)$ *as follows:*

$$q(y_1, \cdots, y_n) = p(x_1(y_1, \cdots, y_n), \cdots, x_n(y_1, \cdots, y_n))|J|;$$

when $(y_1, \cdots, y_n) \in$ *the range domain of* $(f_1, \cdots, f_n)$, *otherwise*, $q(y_1, \cdots, y_n) = 0$.

Proof. Making a change of variables

$$u_1 = f_1(x_1, \cdots, x_n), \cdots, u_n = f_n(x_1, \cdots, x_n)$$

we obtain

$$G(y_1, \cdots, y_n) = \int \cdots \int_D p(x_1, \cdots, x_n) dx_1 \cdots dx_n$$

Proof. Making a change of variables

$$u_1 = f_1(x_1, \cdots, x_n), \cdots, u_n = f_n(x_1, \cdots, x_n)$$

we obtain

$$\begin{aligned} & G(y_1, \cdots, y_n) \\ = & \int \cdots \int_D p(x_1, \cdots, x_n) dx_1 \cdots dx_n \\ = & \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_n} q(u_1, \cdots, u_n) du_1 \cdots du_n. \end{aligned}$$

Hence $q(y_1, \cdots, y_n)$ is the joint density of $(\eta_1, \cdots, \eta_n)$.

Example

If ξ_1 and ξ_2 are independent and uniformly distributed over $(0, 1)$, let

$$\eta_1 = (-2\ln\xi_1)^{1/2}\cos(2\pi\xi_2),$$
$$\eta_2 = (-2\ln\xi_1)^{1/2}\sin(2\pi\xi_2)$$

Then η_1 and η_2 are independent and both follow a normal distribution $N(0, 1)$.

Proof. Let

$$y_1 = (-2\ln x_1)^{1/2}\cos(2\pi x_2),$$
$$y_2 = (-2\ln x_1)^{1/2}\sin(2\pi x_2).$$

Then

$$x_1 = e^{-\frac{y_1^2+y_2^2}{2}}$$
$$x_2 = \frac{1}{2\pi}\text{arcctag}\left(\frac{y_1}{y_2}\right).$$

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} -y_1 e^{-\frac{y_1^2+y_2^2}{2}} & -y_2 e^{-\frac{y_1^2+y_2^2}{2}} \\ \frac{1}{2\pi}\frac{y_2}{y_1^2+y_2^2} & -\frac{1}{2\pi}\frac{y_1}{y_1^2+y_2^2} \end{vmatrix}$$

$$= \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}.$$

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} -y_1 e^{-\frac{y_1^2+y_2^2}{2}} & -y_2 e^{-\frac{y_1^2+y_2^2}{2}} \\ \frac{1}{2\pi}\frac{y_2}{y_1^2+y_2^2} & -\frac{1}{2\pi}\frac{y_1}{y_1^2+y_2^2} \end{vmatrix}$$
$$= \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}.$$

So, the pdf of (η_1, η_2) is

$$q(y_1, y_2) = p(x_1, x_2)|J| = \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y_2^2}{2}}.$$

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} -y_1 e^{-\frac{y_1^2+y_2^2}{2}} & -y_2 e^{-\frac{y_1^2+y_2^2}{2}} \\ \frac{1}{2\pi}\frac{y_2}{y_1^2+y_2^2} & -\frac{1}{2\pi}\frac{y_1}{y_1^2+y_2^2} \end{vmatrix}$$
$$= \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}.$$

So, the pdf of (η_1, η_2) is

$$q(y_1, y_2) = p(x_1, x_2)|J| = \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y_2^2}{2}}.$$

Hence η_1 and η_2 are independent $N(0, 1)$ variables.

Example

Suppose that ξ and η are independent with exponential distributions of parameter 1. Calculate the joint density of $\alpha = \xi + \eta$ and $\beta = \xi/\eta$, and calculate the densities of α, β respectively.

Solution. Observe first that the joint density of (ξ, η) is as follows:

$$p(x, y) = e^{-(x+y)}, \qquad x > 0, y > 0.$$

Also, it is easy to see that $u = x + y, v = x/y \implies x = uv/(1+v), y = u/(1+v)$. When $x, y > 0$, $u, v > 0$ and

$$\begin{aligned} J^{-1} &= \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} 1 & 1 \\ 1/y & -x/y^2 \end{vmatrix} \\ &= -\frac{x+y}{y^2} = -\frac{(1+v)^2}{u}. \end{aligned}$$

Hence we have

$$|J| = \frac{u}{(1+v)^2}.$$

It follows that the joint density of (α, β) is

$$q(u,v) = \begin{cases} \frac{ue^{-u}}{(1+v)^2}, & u > 0, v > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence we have

$$|J| = \frac{u}{(1+v)^2}.$$

It follows that the joint density of (α, β) is

$$q(u, v) = \begin{cases} \frac{ue^{-u}}{(1+v)^2}, & u > 0, v > 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$p_\alpha(u) = ue^{-u}, u > 0, \quad p_\beta(v) = \frac{1}{(1+v)^2}, v > 0.$$

Example

Suppose that ξ and η are i.i.d. with a common normal distribution $N(0,1)$. Let $\rho = \sqrt{\xi^2 + \eta^2}$, $\nu = \xi/\eta$. Prove that ρ and ν are independent.

Proof. By the hypothesis the joint density of (ξ, η) is

Proof. By the hypothesis the joint density of (ξ, η) is

$$p(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right).$$

Proof. By the hypothesis the joint density of (ξ, η) is

$$p(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right).$$

So, the joint distribution of (ρ, ν) is

$$F_{\rho,\nu}(x, y) = P(\rho \leq x, \nu \leq y)$$

Proof. By the hypothesis the joint density of (ξ, η) is

$$p(x,y) = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right).$$

So, the joint distribution of (ρ, ν) is

$$\begin{aligned} F_{\rho,\nu}(x,y) &= P(\rho \le x, \nu \le y) \\ &= P(\sqrt{\xi^2+\eta^2} \le x, \xi/\eta \le y) \end{aligned}$$

Proof. By the hypothesis the joint density of (ξ, η) is

$$p(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right).$$

So, the joint distribution of (ρ, ν) is

$$\begin{aligned} F_{\rho,\nu}(x, y) &= P(\rho \leq x, \nu \leq y) \\ &= P(\sqrt{\xi^2 + \eta^2} \leq x, \xi/\eta \leq y) \\ &= \iint\limits_{\sqrt{u^2+v^2} \leq x, u/v \leq y} \frac{1}{2\pi} \exp(-\frac{u^2 + v^2}{2}) dudv \end{aligned}$$

Letting $u = r\sin\theta$ and $v = r\cos\theta$ yields

$$F_{\rho,\nu}(x,y) = \iint_{0\le r\le x, \tan\theta\le y, -\pi\le\theta<\pi} \frac{1}{2\pi} e^{-r^2/2} r\,dr\,d\theta$$

Letting $u = r\sin\theta$ and $v = r\cos\theta$ yields

$$\begin{aligned} F_{\rho,\nu}(x,y) &= \iint\limits_{0\le r\le x, \tan\theta\le y, -\pi\le\theta<\pi} \frac{1}{2\pi} e^{-r^2/2} r dr d\theta \\ &= \int_0^x e^{-r^2/2} r dr \cdot 2\int_{-\pi/2}^{\tan^{-1} y} \frac{1}{2\pi} d\theta \end{aligned}$$

Letting $u = r\sin\theta$ and $v = r\cos\theta$ yields

$$\begin{aligned} F_{\rho,\nu}(x,y) &= \iint\limits_{0\le r\le x, \tan\theta\le y, -\pi\le\theta<\pi} \frac{1}{2\pi} e^{-r^2/2} r dr d\theta \\ &= \int_0^x e^{-r^2/2} r dr \cdot 2\int_{-\pi/2}^{\tan^{-1} y} \frac{1}{2\pi} d\theta \\ &= (1-e^{-x^2/2}) \cdot \frac{1}{\pi}\Big(\tan^{-1} y + \frac{\pi}{2}\Big), \\ & \quad x>0, -\infty<y<\infty. \end{aligned}$$

The pdf of (ρ, ν) is

$$
\begin{aligned}
f_{\rho,\nu}(x,y) &= \begin{cases} xe^{-x^2/2}\frac{1}{\pi(1+y^2)}, & x>0, -\infty<y<\infty, \\ 0, & \text{otherwise.} \end{cases} \\
&\triangleq f_\rho(x)\cdot f_\nu(y).
\end{aligned}
$$

The pdf of (ρ, ν) is

$$
\begin{aligned}
f_{\rho,\nu}(x,y) &= \begin{cases} xe^{-x^2/2}\frac{1}{\pi(1+y^2)}, & x>0, -\infty<y<\infty, \\ 0, & \text{otherwise.} \end{cases} \\
&\triangleq f_\rho(x)\cdot f_\nu(y).
\end{aligned}
$$

So, ρ and ν are indept.

The pdf of (ρ, ν) is

$$
\begin{aligned}
f_{\rho,\nu}(x,y) &= \begin{cases} xe^{-x^2/2}\frac{1}{\pi(1+y^2)}, & x>0, -\infty<y<\infty, \\ 0, & \text{otherwise.} \end{cases} \\
&\triangleq f_\rho(x)\cdot f_\nu(y).
\end{aligned}
$$

So, ρ and ν are indept.

Here

$$
f_\rho(x) = \begin{cases} xe^{-x^2/2}, & x>0, \\ 0, & \text{otherwise.} \end{cases}
$$

is called Rayleigh distribution.

Example

Suppose $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)' = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{a}$, where $\boldsymbol{C}$ is a $n \times n$ invertible matrix. Find the distribution of $\boldsymbol{\eta}$.

Example

Suppose $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)' = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{a}$, where $\boldsymbol{C}$ is a $n \times n$ invertible matrix. Find the distribution of $\boldsymbol{\eta}$.

Solution. The pdf of $\boldsymbol{\xi}$ is

$$p_{\boldsymbol{\xi}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

Example

Suppose $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)' = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{a}$, where $\boldsymbol{C}$ is a $n \times n$ invertible matrix. Find the distribution of $\boldsymbol{\eta}$.

Solution. The pdf of $\boldsymbol{\xi}$ is

$$p_{\boldsymbol{\xi}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

Let $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{a}$, then $\boldsymbol{x} = \boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})$. It follows that the pdf of $\boldsymbol{\eta}$ is

$$p_{\boldsymbol{\eta}}(\boldsymbol{y}) = p_{\boldsymbol{\xi}}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}))|\boldsymbol{C}^{-1}|$$

$$= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}|\boldsymbol{C}|} \exp\left\{-\frac{1}{2}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})\right\}$$

$$
\begin{aligned}
&= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}|\boldsymbol{C}|} \exp\left\{-\frac{1}{2}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})\right\} \\
&= \frac{1}{(2\pi)^{n/2}|(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{u})'(\boldsymbol{C}^{-1})'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}|\boldsymbol{C}|} \exp\left\{-\frac{1}{2}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})\right\} \\
&= \frac{1}{(2\pi)^{n/2}|(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{u})'(\boldsymbol{C}^{-1})'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})\right\} \\
&= \frac{1}{(2\pi)^{n/2}|(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{C}\boldsymbol{\mu}-\boldsymbol{a})'(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')^{-1}(\boldsymbol{y}-\boldsymbol{C}\boldsymbol{\mu}-\boldsymbol{a})\right\}.
\end{aligned}
$$

$$\begin{aligned}
&= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}|\boldsymbol{C}|}\exp\left\{-\frac{1}{2}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})\right\}\\
&= \frac{1}{(2\pi)^{n/2}|(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{u})'(\boldsymbol{C}^{-1})'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})\right\}\\
&= \frac{1}{(2\pi)^{n/2}|(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{C}\boldsymbol{\mu}-\boldsymbol{a})'(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')^{-1}(\boldsymbol{y}-\boldsymbol{C}\boldsymbol{\mu}-\boldsymbol{a})\right\}.
\end{aligned}$$

So $\boldsymbol{\eta} = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{a} \sim N(\boldsymbol{C}\boldsymbol{\mu} + \boldsymbol{a}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')$.

特别地，如果$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$，
$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)' = \boldsymbol{U}\boldsymbol{\xi}$，$\boldsymbol{U}$ 是正交矩阵，则$\boldsymbol{\eta} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$，
即$N(\mathbf{0}, \sigma^2 \boldsymbol{I})$具有旋转不变性.

特别地，如果$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$,
$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)' = \boldsymbol{U}\boldsymbol{\xi}$, $\boldsymbol{U}$ 是正交矩阵，则$\boldsymbol{\eta} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$,
即$N(\mathbf{0}, \sigma^2 \boldsymbol{I})$具有旋转不变性.

思考题: 反过来，设$\xi_1, \ldots, \xi_n$相互独立，如果对任何正交矩阵$\boldsymbol{U}$,
$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)'$ 与$\boldsymbol{U}\boldsymbol{\xi}$ 同分布，那么$\xi_k \sim N(0, \sigma^2)$, $k = 1, \ldots, n$.

Corollary

If $\boldsymbol{\xi} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *then* $\boldsymbol{\eta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{I})$, *i.e.*, $\eta_1, \ldots, \eta_n$ *are i.i.d. standard normal random variables.*

Corollary

If $\boldsymbol{\xi} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *then* $\boldsymbol{\eta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{I})$, *i.e.,* $\eta_1, \ldots, \eta_n$ *are i.i.d. standard normal random variables.*

Because $\boldsymbol{C} = \boldsymbol{\Sigma}^{-1/2}$, $\boldsymbol{a} = -\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$

$$\boldsymbol{C}\boldsymbol{\mu} + \boldsymbol{a} = \mathbf{0}, \quad \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}' = \boldsymbol{I}.$$

Example

Suppose that X and Y are independent random variables. Assume that the random variable Z depends only on X, and W on Y, that is, $Z = g(X), W = h(Y)$ for g, h, where g and h are Borel functions. Then Z and W are independent.

Proof. For any x and y,

$$P(Z \leq x, W \leq y) = P(g(X) \leq x, h(Y) \leq y)$$

Proof. For any x and y,

$$
\begin{aligned}
& P(Z \le x, W \le y) = P(g(X) \le x, h(Y) \le y) \\
= \; & P\Big(X \in \underbrace{g^{-1}((-\infty, x])}_{B_1 \in \mathcal{B}}, Y \in \underbrace{h^{-1}((-\infty, y]))}_{B_2 \in \mathcal{B}}\Big)
\end{aligned}
$$

Proof. For any x and y,

$$
\begin{aligned}
& P(Z \le x, W \le y) = P(g(X) \le x, h(Y) \le y) \\
= \; & P\Big(X \in \underbrace{g^{-1}((-\infty, x])}_{B_1 \in \mathcal{B}}, Y \in \underbrace{h^{-1}((-\infty, y])}_{B_2 \in \mathcal{B}}\Big) \\
= \; & P\Big(X \in g^{-1}((-\infty, x])\Big) P\Big(Y \in h^{-1}((-\infty, y])\Big) \\
= \; & P(Z \le x) P(W \le y).
\end{aligned}
$$

Proof. For any x and y,

$$
\begin{aligned}
P(Z \le x, W \le y) &= P(g(X) \le x, h(Y) \le y) \\
&= P\Big(X \in \underbrace{g^{-1}((-\infty, x])}_{B_1 \in \mathcal{B}}, Y \in \underbrace{h^{-1}((-\infty, y]))}_{B_2 \in \mathcal{B}}\Big) \\
&= P\Big(X \in g^{-1}((-\infty, x])\Big) P\Big(Y \in h^{-1}((-\infty, y])\Big) \\
&= P(Z \le x) P(W \le y).
\end{aligned}
$$

So, Z and W are indept.

More generally,

Theorem

Let $1 \leq n_1 < n_2 < \cdots < n_k = n$. Assume that f_1 is a Borel function of n_1 arguments,$\cdots$, f_k a Borel function of $n_k - n_{k-1}$ arguments. If $X_1, \cdots, X_n$ are indepet,, then so are $f_1(X_1, \cdots, X_{n_1})$, $f_2(X_{n_1+1}, \cdots, X_{n_2})$, $\cdots$, $f_k(X_{n_{k-1}+1}, \cdots, X_{n_k})$.

In particular, when $f_1, \cdots, f_k$ are functions of a single argument, $f_1(X_1), \cdots, f_k(X_k)$ are indept.

2.5.5 Important distributions in statistics

2.5.5 Important distributions in statistics

χ^2, t and F distributions

χ^2 distribution

χ^2 distribution $\quad \Gamma$ distribution

χ^2 distribution $\quad \Gamma$ distribution

$\xi \sim \Gamma(\lambda, r)$ if it has pdf

$$p(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases} \qquad (\lambda > 0, r > 0)$$

Lemma (Additivity of Gamma distribution) If ξ_1 and ξ_2 are indept., and $\xi_1 \sim \Gamma(\lambda, r_1)$, $\xi_2 \sim \Gamma(\lambda, r_2)$, then $\xi_1 + \xi_2 \sim \Gamma(\lambda, r_1 + r_2)$.

Proof. Let $\eta = \xi_1 + \xi_2$. Obviously, when $z < 0$, $p_\eta(z) = 0$.

Proof. Let $\eta = \xi_1 + \xi_2$. Obviously, when $z < 0$, $p_\eta(z) = 0$. When $z > 0$,

$$p_\eta(z) = \int_o^z p_{\xi_1}(x) p_{\xi_2}(z - x) dx$$

Proof. Let $\eta = \xi_1 + \xi_2$. Obviously, when $z < 0$, $p_\eta(z) = 0$. When $z > 0$,

$$p_\eta(z) = \int_o^z p_{\xi_1}(x) p_{\xi_2}(z - x) dx$$

$$= \int_o^z \frac{\lambda^{r_1}}{\Gamma(r_1)} x^{r_1 - 1} e^{-\lambda x} \frac{\lambda^{r_2}}{\Gamma(r_2)} (z - x)^{r_2 - 1} e^{-\lambda(z - x)} dx$$

Proof. Let $\eta = \xi_1 + \xi_2$. Obviously, when $z < 0$, $p_\eta(z) = 0$. When $z > 0$,

$$
\begin{aligned}
p_\eta(z) &= \int_o^z p_{\xi_1}(x) p_{\xi_2}(z-x)dx \\
&= \int_o^z \frac{\lambda^{r_1}}{\Gamma(r_1)} x^{r_1-1} e^{-\lambda x} \frac{\lambda^{r_2}}{\Gamma(r_2)} (z-x)^{r_2-1} e^{-\lambda(z-x)} dx \\
&\overset{x=zt}{=} \frac{\lambda^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)} z^{r_1+r_2-1} e^{-\lambda z} \int_0^1 t^{r_1-1}(1-t)^{r_2-1}dt
\end{aligned}
$$

Proof. Let $\eta = \xi_1 + \xi_2$. Obviously, when $z < 0$, $p_\eta(z) = 0$. When $z > 0$,

$$\begin{aligned}
p_\eta(z) &= \int_o^z p_{\xi_1}(x) p_{\xi_2}(z-x)dx \\
&= \int_o^z \frac{\lambda^{r_1}}{\Gamma(r_1)} x^{r_1-1} e^{-\lambda x} \frac{\lambda^{r_2}}{\Gamma(r_2)} (z-x)^{r_2-1} e^{-\lambda(z-x)} dx \\
&\overset{x=zt}{=} \frac{\lambda^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)} z^{r_1+r_2-1} e^{-\lambda z} \int_0^1 t^{r_1-1}(1-t)^{r_2-1} dt \\
&= \frac{\lambda^{r_1+r_2}}{\Gamma(r_1+r_2)} z^{r_1+r_2-1} e^{-\lambda z}.
\end{aligned}$$

Therefore, $\eta \sim \Gamma(\lambda, r_1 + r_2)$.

Proof. Let $\eta_1 = \xi_1 + \xi_2$, $\eta_2 = \frac{\xi_1}{\xi_1+\xi_2}$. Then

$$\begin{cases} \xi_1 = & \eta_1\eta_2, \\ \xi_2 = & \eta_1(1-\eta_2). \end{cases} \qquad \begin{cases} x_1 = & y_1y_2, \\ x_2 = & y_1(1-y_2), \end{cases}$$

$y_1 \geq 0, 0 \leq y_2 \leq 1$. Then

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = -y_1.$$

So, the density of (η_1, η_2) is

$$
\begin{aligned}
p(y_1, , y_2) =& \frac{\lambda^{r_1}}{\Gamma(r_1)} (y_1 y_2)^{r_1-1} e^{-\lambda y_1 y_2} \\
&\cdot \frac{\lambda^{r_2}}{\Gamma(r_2)} (y_1(1-y_2))^{r_2-1} e^{-\lambda y_1(1-y_2)} \cdot |y_1| \\
=& \frac{\lambda^{r_1+r_2}}{\Gamma(r_1+r_2)} y_1^{r_1+r_2-1} e^{-\lambda y_1} \\
&\cdot \frac{\Gamma(r_1+r_2)}{\Gamma(r_1)\Gamma(r_2)} y_2^{r_1-1}(1-y_2)^{r_2-1}, \\
& y_1 \geq 0, \quad 0 \leq y_2 \leq 1.
\end{aligned}
$$

So, $\eta_1 = \xi_1 + \xi_2 \sim \Gamma(\lambda, r_1 + r_2)$, $\eta_2 = \frac{\xi_1}{\xi_1 + \xi_2} \sim \beta(r_1, r_2)$.

Example

Suppose that $\xi_1, \cdots, \xi_n$ are independent standard normal random variables. Let

$$\eta = \xi_1^2 + \cdots + \xi_n^2.$$

Find the distribution of η.

Example

Suppose that $\xi_1, \cdots, \xi_n$ are independent standard normal random variables. Let

$$\eta = \xi_1^2 + \cdots + \xi_n^2.$$

Find the distribution of η.

Solution. First, we consider the case of $n = 1$. The cdf of ξ_i^2 is

$$F_{\xi_i^2}(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \phi(u)du, \quad y > 0.$$

Hence the pdf of ξ_i^2 is

$$\begin{aligned} p_{\xi_i^2}(y) =& \phi(\sqrt{y})(\sqrt{y})' - \phi(-\sqrt{y})(-\sqrt{y})' \\ =& \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, \quad y > 0, \end{aligned}$$

which is the pdf of $\Gamma(\frac{1}{2}, \frac{1}{2})$ distribution. So $\xi_i^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$.

Hence the pdf of ξ_i^2 is

$$
\begin{aligned}
p_{\xi_i^2}(y) =&\phi(\sqrt{y})(\sqrt{y})' - \phi(-\sqrt{y})(-\sqrt{y})' \\
=&\frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, \quad y > 0,
\end{aligned}
$$

which is the pdf of $\Gamma(\frac{1}{2}, \frac{1}{2})$ distribution. So $\xi_i^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$.
By the additivity of Gamma distribution,

$$
\eta \sim \Gamma(\frac{1}{2}, \frac{1}{2} + \cdots + \frac{1}{2}) = \Gamma(\frac{1}{2}, \frac{n}{2}).
$$

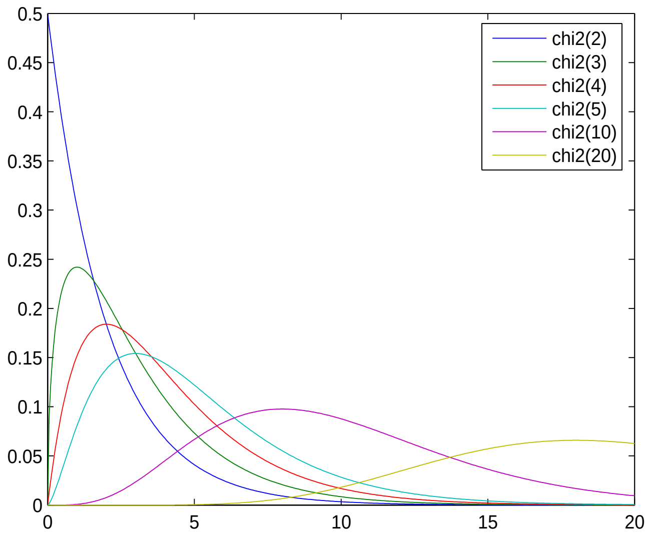
Hence, the pdf of $\eta = \xi_1^2 + \cdots + \xi_n^2$ is

$$p(x) = \begin{cases} \frac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

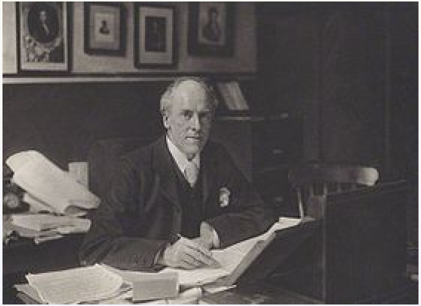
1. The χ^2 distribution

Call $\Gamma(1/2, n/2)$ a $\chi^2(n)$ distribution, where n is the degree of freedom. The density function is

$$p(x) = \begin{cases} \dfrac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x > 0, \\ 0, & x \le 0. \end{cases}$$


chi2(2)
chi2(3)
chi2(4)
chi2(5)
chi2(10)
chi2(20)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0
5
10
15
20

Karl Pearson (March 1857– April 1936)



Theorem

(1) Suppose that $\xi_1, \cdots, \xi_n$ *are independent standard normal random variables, then*

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

(2) The $\chi^2(n)$ *distribution possesses the additivity property. That is, if* $\xi_1 \sim \chi^2(n_1), \xi_2 \sim \chi^2(n_2)$*, and* ξ_1 *and* ξ_2 *are independent, then* $\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$*.*

Theorem

(1) Suppose that $\xi_1, \cdots, \xi_n$ are independent standard normal random variables, then

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

(2) The $\chi^2(n)$ distribution possesses the additivity property. That is, if $\xi_1 \sim \chi^2(n_1), \xi_2 \sim \chi^2(n_2)$, and ξ_1 and ξ_2 are independent, then $\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$.

Proof. (1) had been proved.

Theorem

(1) Suppose that $\xi_1, \cdots, \xi_n$ are independent standard normal random variables, then

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

(2) The $\chi^2(n)$ distribution possesses the additivity property. That is, if $\xi_1 \sim \chi^2(n_1), \xi_2 \sim \chi^2(n_2)$, and ξ_1 and ξ_2 are independent, then $\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$.

Proof. (1) had been proved. (2) follows from the additivity of Gamma distribution immediately.

Corollary

If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *then* $(\boldsymbol{\xi} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \sim \chi^2(n)$.

Corollary

If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *then* $(\boldsymbol{\xi} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \sim \chi^2(n)$.

Proof. Let $\boldsymbol{\eta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu})$. Then $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{I})$. That is, $\eta_1, \cdots, \eta_n$ are i.i.d. standard normal random variables.

Corollary

If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *then* $(\boldsymbol{\xi} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu}) \sim \chi^2(n)$.

Proof. Let $\boldsymbol{\eta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu})$. Then $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{I})$. That is, $\eta_1, \cdots, \eta_n$ are i.i.d. standard normal random variables. So

$$\begin{aligned}&(\boldsymbol{\xi} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu}) = \boldsymbol{\eta}'\boldsymbol{\eta}\\ =&\eta_1^2 + \cdots + \eta_n^2 \sim \chi^2(n).\end{aligned}$$

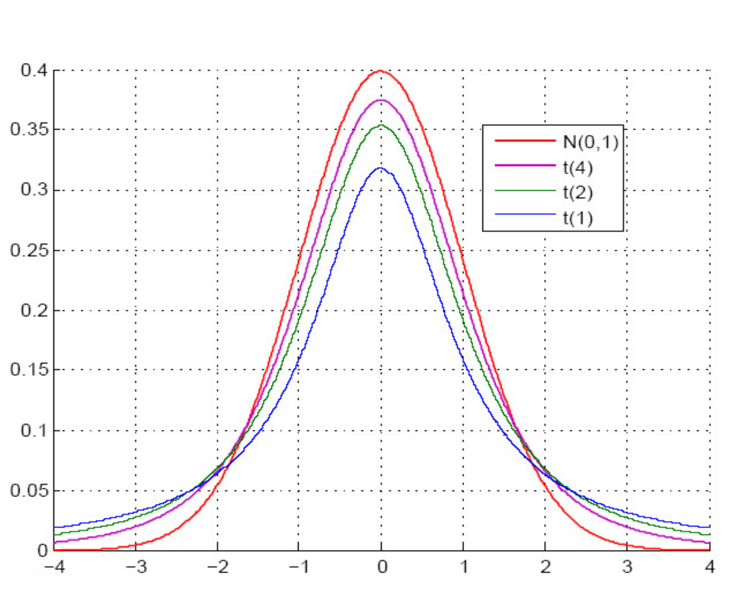
2. The t-distribution

Theorem

If ξ and η are independent, and $\xi \sim N(0,1), \eta \sim \chi^2(n)$, then the random variable $T = \frac{\xi}{\sqrt{\eta/n}}$ *has the density*

$$p(t) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\,\Gamma(n/2)}(1+t^2/n)^{-(n+1)/2}, \quad -\infty < t < \infty.$$

We call the random variable T above a $t(n)$ distribution with n as its degree of freedom.


N(0,1)
t(4)
t(2)
t(1)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−4
−3
−2
−1
0
1
2
3
4

William Gosset(1876–1937)

- 1908年提出t-分布

证明：令$S=\eta$. 考察变换:

$$\begin{cases} t = \frac{x}{\sqrt{y/n}}, \\ s = y; \end{cases} \qquad \begin{cases} x = t\sqrt{s/n}, \\ y = s. \end{cases}$$

则

$$J = \frac{\partial(x,y)}{\partial(t,s)} = \begin{vmatrix} \sqrt{s/n} & \frac{t\sqrt{1/n}}{2\sqrt{s}} \\ 0 & 1 \end{vmatrix} = \sqrt{s/n}.$$

所以(T,S)的密度函数为

$$\begin{aligned} p(t,s) =& \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2 s/n}{2}} \frac{(1/2)^{n/2}}{\Gamma(n/2)} s^{n/2-1} e^{-s/2} \sqrt{s/n} \\ =& \frac{(1/2)^{(n+1)/2}}{\sqrt{n\pi}\Gamma(n/2)} s^{\frac{n+1}{2}-1} \exp\left\{-s\left(\frac{t^2}{2n}+\frac{1}{2}\right)\right\}, \\ & -\infty < t < \infty, \quad s \geq 0. \end{aligned}$$

所以(T,S)的密度函数为

$$\begin{aligned}
p(t,s) =& \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2 s/n}{2}} \frac{(1/2)^{n/2}}{\Gamma(n/2)} s^{n/2-1} e^{-s/2} \sqrt{s/n} \\
=& \frac{(1/2)^{(n+1)/2}}{\sqrt{n\pi}\Gamma(n/2)} s^{\frac{n+1}{2}-1} \exp\left\{-s\left(\frac{t^2}{2n}+\frac{1}{2}\right)\right\}, \\
& -\infty < t < \infty, \quad s \geq 0.
\end{aligned}$$

因此T的密度函数为

$$\begin{aligned}
p(t) =& \int_0^\infty p(t,s)ds = \frac{(1/2)^{(n+1)/2}\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(n/2)} \left(\frac{t^2}{2n}+\frac{1}{2}\right)^{-\frac{n+1}{2}} \\
=& \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(n/2)} \left(1+t^2/n\right)^{-\frac{n+1}{2}}, \quad -\infty < t < \infty.
\end{aligned}$$

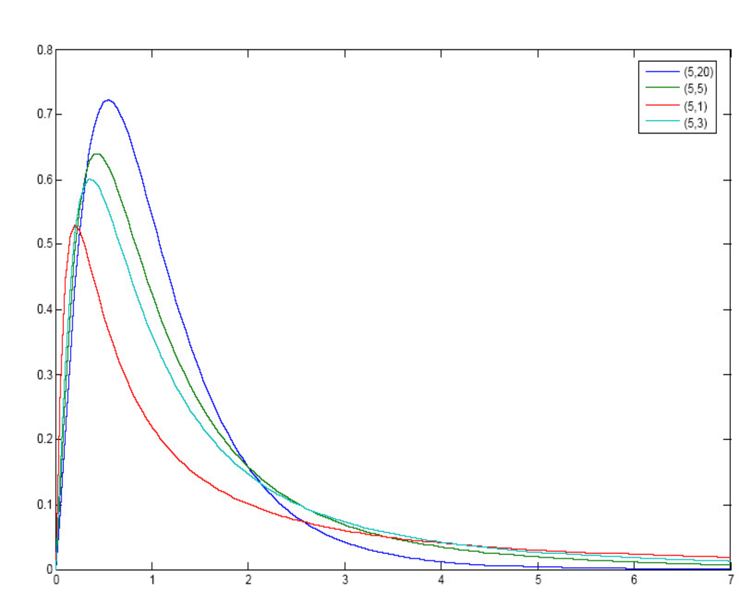
3. The F-distribution Fisher – Snedecor distribution (after Ronald Fisher and George W. Snedecor)

Theorem

Suppose that ξ and η are independent, and $\xi \sim \chi^2(m), \eta \sim \chi^2(n)$, then the random variable $F = \frac{\xi/m}{\eta/n}$ has the density

$$p(x) = \begin{cases} \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} m^{m/2} n^{n/2} \frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}}, & x > 0, \\ 0, & x \le 0 \end{cases}$$

We call the random variable F above an $F(m,n)$ distribution with m and n as its first and second degrees of freedom respectively.


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
(5,20)
(5,5)
(5,1)
(5,3)

Ronald Aylmer Fisher (February 1890 – July 1962)

George Waddel Snedecor (October 1881 – February 1974)

Figure: R.A. Fisher

证明： 令$S=\eta$. 考察变换:

$$\begin{cases} t=\dfrac{x/m}{y/n}, \\ s=y; \end{cases} \qquad \begin{cases} x=\dfrac{m}{n}ts, \\ y=s. \end{cases}$$

则

$$J=\frac{\partial(x,y)}{\partial(t,s)}=\begin{vmatrix} \frac{m}{n}s & \frac{m}{n}t \\ 0 & 1 \end{vmatrix}=\frac{m}{n}s.$$

所以(F,S)的密度函数为

$$
\begin{aligned}
p(t,s) =& \frac{\left(\frac{1}{2}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)}\left(\frac{m}{n}ts\right)^{\frac{m}{2}-1} e^{-\frac{m}{2n}ts} \cdot \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} s^{\frac{n}{2}-1} e^{-s/2} \cdot \frac{m}{n} s \\
=& \frac{\left(\frac{1}{2}\right)^{\frac{m+n}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} t^{\frac{m}{2}-1} s^{\frac{m+n}{2}-1} \exp\left\{-s\left(\frac{m}{n}t+1\right)\frac{1}{2}\right\}, \\
& t, s \geq 0.
\end{aligned}
$$

因此F的密度函数为

$$
\begin{aligned}
p(t) =& \int_0^\infty p(t,s)ds = \frac{\left(\frac{1}{2}\right)^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} t^{\frac{m}{2}-1}\left(\left(\frac{m}{n}t+1\right)\frac{1}{2}\right)^{-\frac{m+n}{2}} \\
=& \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{t^{\frac{m}{2}-1}}{(mt+n)^{\frac{m+n}{2}}}, \quad t \geq 0.
\end{aligned}
$$

The F-distribution possesses the following properties:

(1) If $F \sim F(m,n)$, then $1/F \sim F(n,m)$.

(2) If $T \sim t(n)$, then $T^2 \sim F(1,n)$.

The F-distribution possesses the following properties:

(1) If $F \sim F(m,n)$, then $1/F \sim F(n,m)$.

(2) If $T \sim t(n)$, then $T^2 \sim F(1,n)$.

Proof. (1) Simple. It immediately follows from the definition of F.

The F-distribution possesses the following properties:
(1) If $F \sim F(m,n)$, then $1/F \sim F(n,m)$.
(2) If $T \sim t(n)$, then $T^2 \sim F(1,n)$.
Proof. (1) Simple. It immediately follows from the definition of F. (2) Write $T = \xi/\sqrt{\eta/n}$, where ξ and η are independent and $\xi \sim N(0,1), \eta \sim \chi^2(n)$. Note that $T^2 = \xi^2/(\eta/n)$. Also, $\xi^2 \sim \chi^2(1)$ and ξ^2, η are independent. Hence $T^2 \sim F(1,n)$.

4. Simulating the distribution

In many cases, the analytic formula of the cdf of $Y = f(X_1, \cdots, X_n)$ is difficult (or impossible) to derive, though the cdf of $\boldsymbol{X} = (X_1, \cdots, X_n)'$ is known. In some case, the cdf of Y is too complex for applications. For example,

$$T = \max_{0 \le i,j \le k} |X_i - X_j|,$$

where $X_i \sim N(0, 1/n_i)$, $i = 1, 2, \cdots, k$, are indept.

The cdf of T is important in statistics. But the analytic formula of its cdf is very complex.

In statistics, there is a method to obtain the approximation of the cdf.

In statistics, there is a method to obtain the approximation of the cdf.

Notice

$$F_Y(x) = P(A), \quad A = \{Y \le x\}.$$

If we can repeat a trial related to A a lot of times, then

$$F_Y(x) = P(A) \approx \text{frequency of } A.$$

In statistics, there is a method to obtain the approximation of the cdf.

Notice

$$F_Y(x) = P(A), \quad A = \{Y \leq x\}.$$

If we can repeat a trial related to A a lot of times, then

$$F_Y(x) = P(A) \approx \text{frequency of } A.$$

Simulation or Monte Carlo method

- Step 1, using the cdf of $\boldsymbol{X} = (X_1, \cdots, X_n)'$, generate a random number $\boldsymbol{X} = \boldsymbol{x}$;

- Step 1, using the cdf of $\boldsymbol{X} = (X_1, \cdots, X_n)'$, generate a random number $\boldsymbol{X} = \boldsymbol{x}$;
- Step 2, compute the value of $Y = f(\boldsymbol{X}) = f(\boldsymbol{x})$ denoted by y_1;

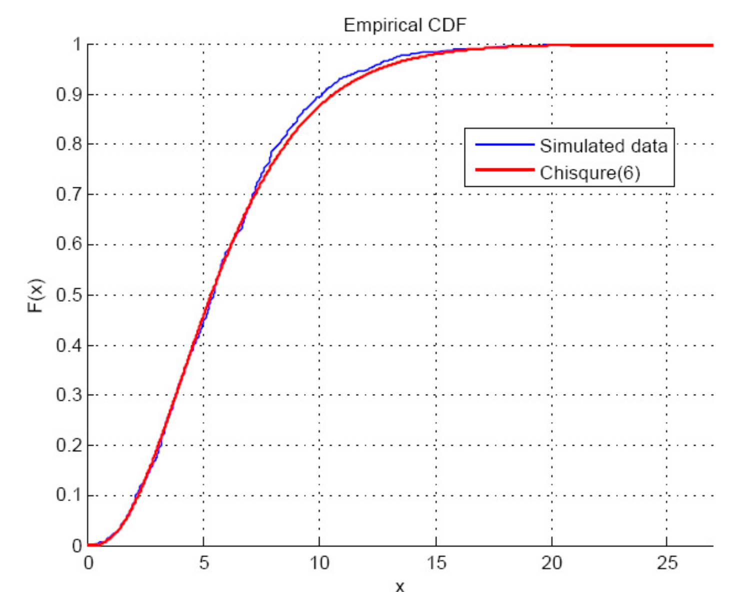
- Step 1, using the cdf of $\boldsymbol{X} = (X_1, \cdots, X_n)'$, generate a random number $\boldsymbol{X} = \boldsymbol{x}$;
- Step 2, compute the value of $Y = f(\boldsymbol{X}) = f(\boldsymbol{x})$ denoted by y_1;
- Step 3, repeat Steps 1-2 N times $(N = 10,000, N = 100,100, N = 1,000,000)$, obtain $y_1, \cdots, y_N$;

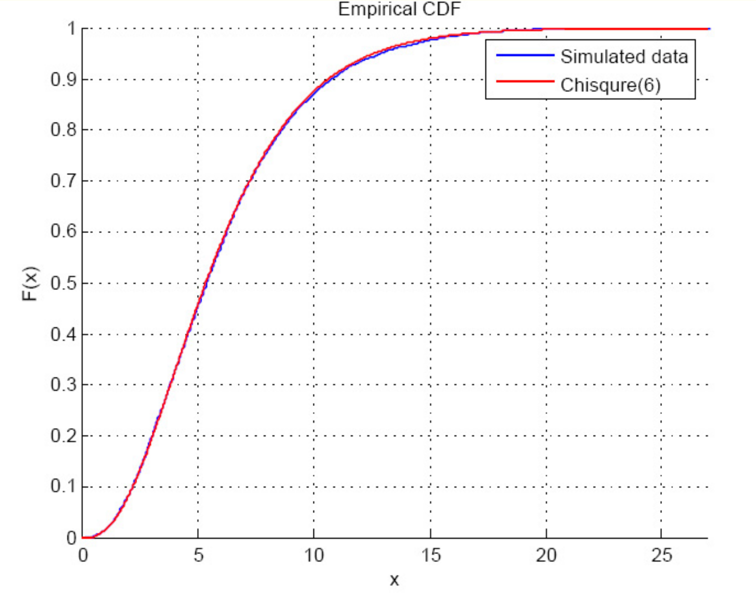
- Step 1, using the cdf of $\boldsymbol{X}=(X_1,\cdots,X_n)'$, generate a random number $\boldsymbol{X}=\boldsymbol{x}$;
- Step 2, compute the value of $Y=f(\boldsymbol{X})=f(\boldsymbol{x})$ denoted by y_1;
- Step 3, repeat Steps 1-2 N times ($N=10,000, N=100,100, N=1,000,000$), obtain $y_1,\cdots,y_N$;
- Step 4,

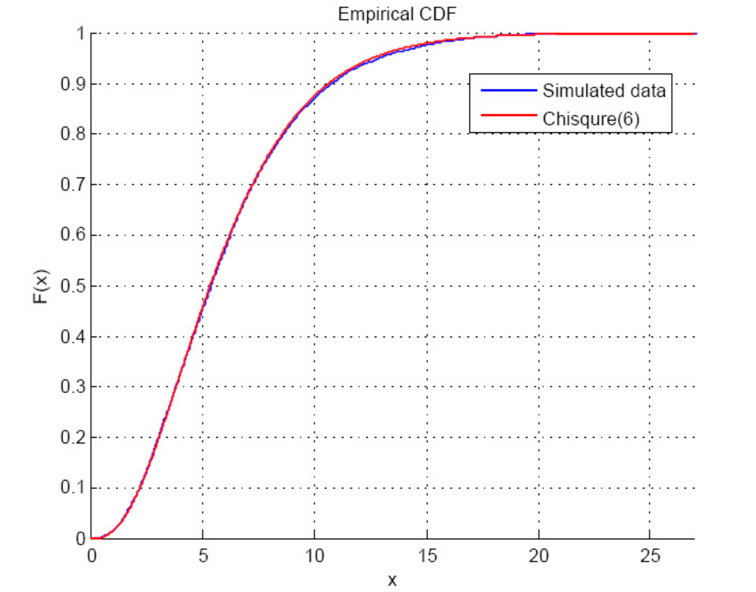
$$F_{\boldsymbol{Y}}(y)\approx F_N(y)=\frac{\#\{i: y_i\le y\}}{N}.$$

Example

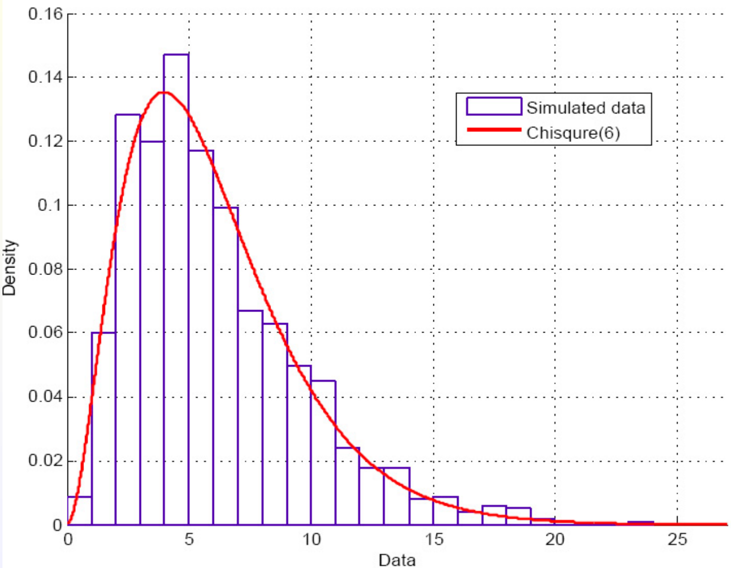
$\chi^2 = \xi_1^2 + \cdots + \xi_6^2$, $\xi_1, \cdots, \xi_6$ i.i.d. $\sim N(0,1)$. $N = 1,000,000$.

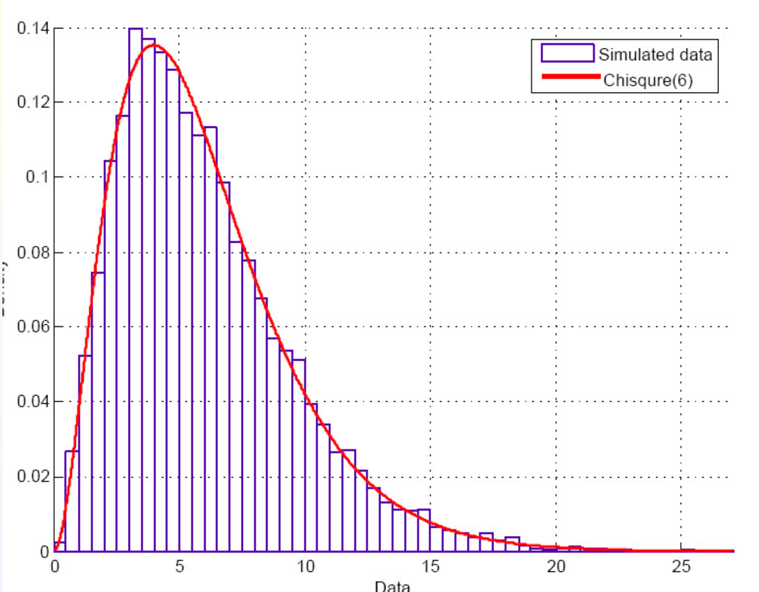

Empirical CDF
F(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25
Simulated data
Chisqure(6)

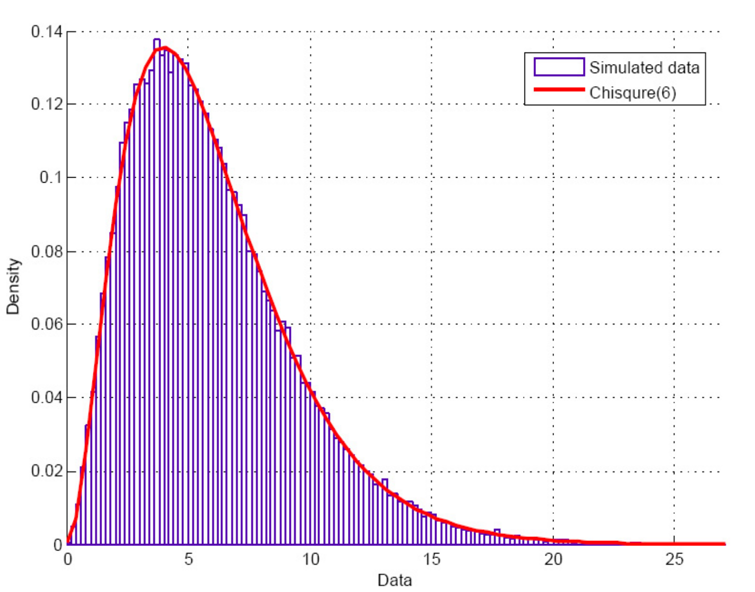

Empirical CDF
Simulated data
Chisqure(6)
F(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25


Empirical CDF
Simulated data
Chisqure(6)
F(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25

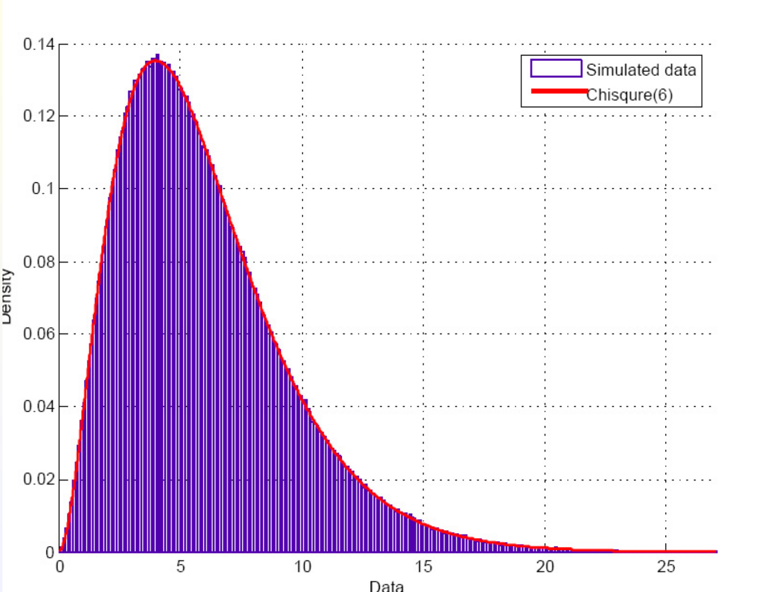
Simulation pdf-figs $N = 1,000$




0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
Data
Simulated data
Chisqure(6)


0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Density
0
5
10
15
20
25
Data
Simulated data
Chisqure(6)

Simulation pdf-figs $N = 1,000,000$



设$f(x)$, $g(y)$ 为密度函数, $g(y)>0$. 并且存在常数$c>0$满足

$$\frac{f(y)}{g(y)} \le c, \quad \forall\, y.$$

现设$Y_1, U_1, Y_2, U_2, \cdots$, 为一列独立随机变量, Y_i的密度函数都为$g(y)$, U_i都为$[0,1]$上的均匀随机变量.
定义X如下: 若$U_1 \le \frac{f(Y_1)}{cg(Y_1)}$, 则令$X = Y_1$, 否则再考虑$U_2, Y_2$, 若$U_2 \le \frac{f(Y_2)}{cg(Y_2)}$, 则令$X = Y_2$, 否则再考虑$U_3, Y_3$, 以此类推.
证明: X的密度函数为$f(y)$.