Nonlinear Analysis and Differential Equations, Vol. 5, 2017, no. 5, 209 - 216
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/nade.2017.756

Properties of the Random Variable in Normal Distribution

Jian Wang and Jian Xin Wang

School of Science
Shandong University of Technology
Zibo, China

Abstract

In this paper, we proved some new properties of normal distribution random variables, and these properties will provide better understanding of the foundation of normal distribution.

Keywords: normal distribution, random variable, properties

Introduction

Let random variable X obey a probability distributions of the location parameter μ and the scale parameter σ. The probability density function is $f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$. Then the random variable calls normal random variable, i.e. $X \sim N(\mu, \sigma^2)$.

Property 1: Let X and Y are random variables. If we know that $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$, X and Y are independent of one another. Hence,

(1)$X + Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ (2)$X - Y \sim N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$
(3)$X - Y$ and $X + Y$ are independent of one another if $\sigma_1^2 = \sigma_2^2$

Proof:

Since

$$X \sim N(\mu_1, \sigma_1^2) \ , Y \sim N(\mu_2, \sigma_2^2)$$

it is easily verified that the characteristic functions are

$$f_X(t) = \exp\left(i\mu_1 t - \frac{1}{2}\sigma_1^2 t^2\right), f_Y(t) = \exp\left(i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2\right)$$

Hence

$$f_{X+Y}(t) = f_X(t) f_Y(t) = \exp\{i(\mu_1 + \mu_2)t - \tfrac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2\} \tag{a}$$

Besides, there is a one-to-one correspondence between the characteristic function and the distribution function, and (a) is a characteristic function of $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. Therefore, we can show that $X + Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

(2) Since

$$f_Y(t) = \exp\left(i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2\right)$$

Let $Z = -Y$, we have

$$f_Z(t) = f_{-Y}(t) = f_Y(-t) = \exp\left(-i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2\right)$$

Besides, $X - Y = X + Z$.

Hence

$$f_{X-Y}(t) = f_{X+Z}(t) = f_X(t) f_Z(t) = \exp\{i(\mu_1 - \mu_2)t - \tfrac{1}{2}(\sigma_1^2+\sigma_2^2)t^2\} \tag{b}$$

Besides, there is a one-to-one correspondence between the characteristic function and the distribution function, and (b) is a characteristic function of $N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$. Therefore, we can show that $X - Y \sim N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$.

(3) By the above results, we obtain that $X - Y$ and $X + Y$ are univariate normal distribution. Hence, $X - Y$ and $X + Y$ independence is equivalent to uncorrelated. Since

$$E(X^2) = \sigma_1^2 + \mu_1^2 \qquad E(Y^2) = \sigma_2^2 + \mu_2^2$$
$$E(X + Y) = \mu_1 + \mu_2 \qquad E(X - Y) = \mu_1 - \mu_2$$

we have

$$Cov(X + Y, X - Y) = E\big((X + Y)(X - Y)\big) - E(X + Y)E(X - Y)$$
$$= E(X^2 - Y^2) - (\mu_1 + \mu_2)(\mu_1 - \mu_2) = \sigma_1^2 + \mu_1^2 - \sigma_2^2 - \mu_2^2 - \mu_1^2 + \mu_2^2$$
$$= \sigma_1^2 - \sigma_2^2$$

Therefore, $X - Y$ and $X + Y$ are independent of one another when and only when $\sigma_1^2 = \sigma_2^2$.

Property 2: Let X and Y be random variables, and they are independent of each other. The random variables X and Y are normal distributions if $S = X + Y$ is normal distribution.

Proof: By hypothesis, $S \sim N(\mu, \sigma^2)$. We have the following characteristic function

$$f_s(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2} = f_{X+Y}(t) = f_X(t) f_Y(t) \underset{\sigma^2 = \sigma_1^2 + \sigma_2^2}{\overset{\mu = \mu_1 + \mu_2}{=}} e^{i\mu_1 t - \frac{1}{2}\sigma_1^2 t^2} \cdot e^{i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2}$$

Hence

$$f_X(t) = e^{i\mu_1 t - \frac{1}{2}\sigma_1^2 t^2} \qquad f_Y(t) = e^{i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2}$$

From the uniqueness theorem of characteristic function, we can come to the conclusion that the random variables X and Y are also normal distributions.
Moreover, this property can be extended to other distributions. For example, Poisson distribution, binomial distribution and multinomial distribution.

Property 3: Let $Z_1 = X_1 + X_2$, $Z_2 = X_1 - X_2$. The random variables X_1 and X_2 are also normal distributions if Z_1 and Z_2 are normal distributions and they are independent of each other.

Proof : Since

$$Z_1 = X_1 + X_2, Z_2 = X_1 - X_2$$

we obtain

$$X_1 = \frac{Z_1 + Z_2}{2} \quad \text{and} \quad X_2 = \frac{Z_1 - Z_2}{2}$$

Because Z_1 and Z_2 are normal distributions and they are independent of each other. Hence, $Z_1 + Z_2$ and $Z_1 - Z_2$ are also normal distributions.
Besides, X_1 and X_2 are the linear combination of $Z_1 + Z_2$ and $Z_1 - Z_2$. By the linearity of normal distribution, we easily obtain that the random variables X_1 and X_2 are also normal distributions.

Property 4: Let X and Y be random variables and they are independent identically distributed. Their density functions are not equal to zero, and they have second derivatives. We can show that $X, Y, X + Y$ and $X - Y$ are normal distributions if $X + Y$ and $X - Y$ are independent of each other.

Proof: Let $p(x)$ be density function of the random variable X, and $p(y)$ be density function of the random variable Y. The density functions of $U = X + Y$ and $V = X - Y$ are $f(u)$ and $g(v)$, respectively. Besides, U and V are independent of each other. Hence

$$\begin{cases} U = X + Y \\ V = X - Y \end{cases} \Rightarrow \begin{cases} u = x + y \\ v = x - y \end{cases} \Rightarrow \begin{cases} x = \dfrac{u+v}{2} \\ y = \dfrac{u-v}{2} \end{cases}$$

Then $$p\left(\frac{u+v}{2}\right) p\left(\frac{u-v}{2}\right) \cdot \frac{1}{2} = f(u)g(v) \tag{a}$$

We need to take log of (a)

$$\ln p\left(\frac{u+v}{2}\right) + \ln p\left(\frac{u-v}{2}\right) - \ln 2 = \ln f(u) + \ln g(v) \tag{b}$$

Let $h(z) = \ln p(z)$. Then

$$h\left(\frac{u+v}{2}\right) + h\left(\frac{u-v}{2}\right) - \ln 2 = \ln f(u) + \ln g(v) \tag{c}$$

We take partial derivative to u of both sides of (c). We have

$$\frac{1}{2}h'\left(\frac{u+v}{2}\right) + \frac{1}{2}h'\left(\frac{u-v}{2}\right) = \frac{f'(u)}{f(u)} \tag{d}$$

Next, we take partial derivative to v both sides of (d). We have

$$\frac{1}{4}h''\left(\frac{u+v}{2}\right) - \frac{1}{4}h''\left(\frac{u-v}{2}\right) = 0 \tag{e}$$

Let $u = v$. Hence, $h''(u) - h''(0) = 0$. Let $h''(0) = -\lambda$, and we take u rewrite as x. Then

$$\ln p(x) = -\frac{\lambda}{2}x^2 + c_1 x + c_2 \tag{f}$$

Therefore $p(x) = e^{-\frac{\lambda}{2}x^2 + c_1 x + c_2} = e^{c_2 + \frac{c_1^2}{2\lambda}} \cdot e^{-\frac{\lambda}{2}\left(x - \frac{c_1}{\lambda}\right)^2} = ke^{-\frac{\lambda}{2}\left(x - \frac{c_1}{\lambda}\right)^2}$

We know that $\lambda > 0$ ($p(x)$ is a density function), and let $\mu = \frac{c_1}{\lambda}$

Since $\int_{-\infty}^{\infty} p(x)dx = 0$, then $k = \sqrt{\frac{\lambda}{2\pi}}$.

Hence

$$p(x) = \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{(x-\mu)^2}{2\frac{1}{\lambda}}}$$

It is the density function of $N(\mu, \frac{1}{\lambda})$ and, let $x = 0$ in (f), then $c_2 = \ln p(0)$.

We take the derivative to x for (f). We have $\frac{p'(x)}{p(x)} = -\lambda x + c_1$, and $c_1 = \frac{p'(0)}{p(0)}$ (when $x = 0$).
Therefore, we can show that $X, Y, X+Y$ and $X-Y$ are normal distributions if $X+Y$ and $X-Y$ are independent of each other.

Property 5: Let U_1 and U_2 be random variables, and they are independent of each other. They all obey $U(0,1)$. Then the random variables ξ and η are independent of each other and obey $N(0,1)$ if $\xi = (-2\ln U_1)^{\frac{1}{2}}\cos 2\pi U_2$ as well as $\eta = (-2\ln U_1)^{\frac{1}{2}}\sin 2\pi U_2$.

Proof: It is easily verified that the joint probability density function of U_1 and U_2 is $p(u_1, u_2) = 1$.
Then

$$\begin{cases} \xi = (-2\ln U_1)^{\frac{1}{2}}\cos 2\pi U_2 \\ \eta = (-2\ln U_1)^{\frac{1}{2}}\sin 2\pi U_2 \end{cases} \Rightarrow \begin{cases} x = (-2\ln u_1)^{\frac{1}{2}}\cos 2\pi u_2 \\ y = (-2\ln u_1)^{\frac{1}{2}}\sin 2\pi u_2 \end{cases}$$

$$\Rightarrow \begin{cases} u_1 = e^{-\frac{x^2+y^2}{2}} \\ u_2 = \dfrac{1}{2\pi}\arcsin\dfrac{y}{\sqrt{x^2+y^2}} \end{cases}$$

The Jacobi determinant is

$$J = \begin{vmatrix} \dfrac{\partial u_1}{\partial x} & \dfrac{\partial u_1}{\partial y} \\ \dfrac{\partial u_2}{\partial x} & \dfrac{\partial u_2}{\partial y} \end{vmatrix} = -\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$$

Therefore, the joint probability density function of (ξ, η) is

$$p(x,y) = 1 \cdot |J| = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$$

Hence, the random variables ξ and η are independent of each other, and they obey $N(0,1)$. Besides, we can generate random numbers of standard normal distribution by using this property, and generate any parameters random numbers of normal distribution.

Property 6: Normal distribution entropy is the largest when given the mean and variance in a probability distribution.
(In information theory, entropy was introduced by Claude Elwood Shannon, the

founder of information theory, and it was used to measure the degree of uncertainty of experiment. We can defining the entropy by

$$H(\alpha) = -\int_{-\infty}^{\infty} p(x)\log p(x)\,dx$$

For convenience, we use the natural logarithm.)

Let $p(x)$ be univariate density function, with mean μ and variance σ^2. Then the probability distribution of maximum entropy is normal distribution.

Proof: When have maximum entropy, we know that $\partial H = \frac{\partial H(\alpha)}{\partial p(x)} = 0$.

Constraint conditions are as follows:

(1) $P = \int_{-\infty}^{\infty} p(x)dx = 1$ (2) $\mu = \int_{-\infty}^{\infty} xp(x)dx = 1$

(3) $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$

Based on the variational principle and three constraint conditions, we introduce three undetermined multipliers λ, α and β.

Then

$$\partial H + \lambda P + \alpha\mu + \beta\sigma^2 = 0$$

Hence

$$\partial \int_{-\infty}^{\infty} [p(x)\log p(x) - \lambda p(x) - \alpha x p(x) - \beta(x-\mu)^2 p(x)]dx = 0$$

The integrand is $F = p(x)\log p(x) - \lambda p(x) - \alpha x p(x) - \beta(x-\mu)^2 p(x)$ and it is a function of $p(x)\ and\ x$.

Therefore

$$\frac{\partial F}{\partial p} = \log p(x) - \alpha x - \beta(x-\mu)^2 - (\lambda - 1) = 0$$

and

$$p(x) = e^{\lambda - 1} e^{\alpha x + \beta(x-\mu)^2} = Ce^{\beta\left[x - \left(\mu - \frac{\alpha}{2\beta}\right)\right]^2}$$

Here we replaced C with the modification of λ, and it was undetermined coefficient.

We have used the constraint condition (1). Let $y = x - \left(\mu - \frac{\alpha}{2\beta}\right)$, because of the integral of the function is 1, and $\beta < 0$. Then

$$\int_{-\infty}^{\infty} p(x)dx = \int_{-\infty}^{\infty} p(y)dy = \int_{-\infty}^{\infty} Ce^{\beta y^2}dy = C\sqrt{\frac{-\pi}{\beta}} = 1$$

and $C = \sqrt{\frac{-\beta}{\pi}}$. Using the constraint condition (2), we have

$$\int_{-\infty}^{\infty} xp(x)dx = \int_{-\infty}^{\infty} \left(y + \mu - \frac{\alpha}{2\beta}\right) p(y)dy = \mu$$

We note that $\int_{-\infty}^{\infty} yp(y)dy = \frac{\alpha}{2\beta}$. Besides, $yp(y)$ is odd function. Hence, $\alpha = 0, y = x - \mu$ and $p(x) = \sqrt{\frac{-\beta}{\pi}} e^{\beta(x-\mu)^2}$.

Using the constraint condition (3), we have

$$\int_{-\infty}^{\infty} y^2 p(y)dy = \sqrt{-\frac{\beta}{\pi}} \frac{\sqrt{\pi}}{2(-\beta)^{\frac{3}{2}}} = -\frac{1}{2\beta} = \sigma^2$$

We note that $\beta = -\frac{1}{2\sigma^2}$.

Therefore, the density function of normal distribution is

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Consider the second variation:

$$\frac{\partial^2 F}{\partial(p(x))^2} = \frac{1}{p(x)} > 0, \partial^2 H < 0$$

i.e. normal distribution achieve the maximum of information entropy.

At the moment,

$$H(\alpha) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left[\log\sqrt{2\pi}\sigma + \frac{x^2}{2\sigma^2}\right] dx$$

$$= \log\sqrt{2\pi}\sigma + \frac{1}{2} = \log\sqrt{2\pi e}\sigma$$

we have found that the entropy size of normal distribution depends on variance size.

References

[1] L.Q. Xiong, Some Properties on Normal Distribution, *Mathematical Theory and Application*, **20** (2000), 103-105.

[2] Variational method and multiplier method: Normal distribution is derived by the maximum entropy, http://tieba.baidu.com/p/1481901460. 2012

[3] X.P. Li and Z.Y. Chen, *Probability Theory based Learning Instruction*, Higher Education Press, 2011.

[4] Z.F. Wang, Z.X. Su and J. Guan, Important conclusion of and notes to normal distribution, *J. Bohai University*, **32** (2011), 97-99.

Received: May 15, 2017; Published: June 15, 2017