Basic of Distribution

listenzcc

July 21, 2020

Abstract

A family of *normal* distributions, like Normal, Chi-squared and Student's t-distribution. The Normal Distribution is the *core* conception, all others are derived from it.

This article begins with Gamma and Beta function. Since they are useful for computing the *moments* the normal distribution family.

Then the distributions are described one by one.

The Appendix provides some necessary proofs.

Contents

1 Pre-knowledge

1.1 Gamma and Beta function

An infinity integral is called as *Gamma (Γ) function*

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

The *Beta (B) function* is a two-factor function, derived from Γ function

$$B(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)} \tag{2}$$

1.2 Important equations

Proposition 1.1. *Some very important equations.*

The value of $\Gamma(\frac{1}{2})$

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

The recursive of $\Gamma(n)$*, the general situation,*

$$\Gamma(1+z) = z\Gamma(z)$$
$$\Gamma(1-z) = -z\Gamma(-z)$$

The integer situation,

$$\Gamma(n) = (n-1)! \quad \forall n \in \mathcal{N}^+$$

The relationship between Γ *and* e^{-x^2}

$$\Gamma(z) = 2\int_0^\infty x^{2z-1}e^{-x^2}\,\mathrm{d}x$$

The relationship between Γ *and* B *Function*

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t$$

See A.1 for proof.

2 Normal distribution

2.1 Definition

It is hard to say normal distribution is what, since almost every thing follows it.

The Probability Distribution Function *(pdf)* of normal distribution is

$$p(x) = \frac{1}{\sqrt{2\pi}\delta}\exp(-\frac{(x-\mu)^2}{2\delta^2}), -\infty < x < \infty \tag{3}$$

the symbolic notion is $p(x) \sim N(\mu, \delta^2)$. When $\mu = 0$ and $\delta^2 = 1$, it is called standard normal distribution.

2.2 Mean and Variance

The mean and variance of the normal distribution is

$$Mean \triangleq E(x) = \mu$$
$$Variance \triangleq E(x^2) - E^2(x) = \delta^2$$

it is easy to proof using Proposition 1.1.

3 Chi-squared distribution

3.1 Definition

If $Y_i \sim N(0,1)$, then

$$\chi^2 \equiv \sum_{i=1}^{r} Y_i^2 \tag{4}$$

is distributed as Chi-squared χ^2 distribution with r degrees of freedom. The symbolic notion is $p_r(x) \sim \chi^2(r)$.

The pdf of Chi-squared distribution is

$$P_r(x) = \frac{x^{r/2-1}e^{-x/2}}{\Gamma(r/2)2^{r/2}}, 0 < x < \infty \tag{5}$$

3.2 Mean and Variance

The mean and variance of the chi-squared distribution is

$$Mean \triangleq E(x) = r$$
$$Variance \triangleq E(x^2) - E^2(x) = 2r$$

it is easy to proof using Proposition 1.1.

3.3 Relationship with Normal Distribution

The Chi-squared distribution is derived from Normal Distribution.

4 Student's t-distribution

4.1 Definition

The probability distribution of a random variable T, of the form

$$T = \frac{\bar{x} - m}{s/\sqrt{N}}$$

where $\bar{x}$ is the sample mean value of all N samples, m is the population mean value and s is the population standard deviation.

Or, in a more formal one

$$T = \frac{X}{\sqrt{Y/r}} \tag{6}$$

where $X \sim N(0,1)$ and $Y \sim \chi_r^2$.

The pdf of Student's t-distribution is

$$t_r(x) = \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})\sqrt{r\pi}}(1 + \frac{x^2}{r})^{-\frac{r+1}{2}}, -\infty < x < \infty \tag{7}$$

it is easy to proof the pdf is a pdf using A.2.

The pdf of Student's t-distribution can be computed using A.3.

4.2 Relationship with Normal Distribution

It is easy to see that $\lim_{r\to\infty} t_r(x) \sim N(0,1)$. It demonstrates that when r is large enough, the Student's t-distribution is equalize to Normal Distribution.

4.3 Mean and Variance

The mean and variance of the Student's t-distribution is

$$Mean \triangleq E(x) = 0$$
$$Variance \triangleq E(x^2) - E^2(x) = \frac{r}{r-2}$$

A Appendix

A.1 The relationship between Γ and $B(\alpha, \beta)$

It is essential to know that

$$\Gamma(m)\Gamma(n) = B(m,n)\Gamma(m+n)$$

Proof. One can write

$$\Gamma(m)\Gamma(n) = \int_0^\infty x^{m-1}e^{-x}dx \int_0^\infty y^{n-1}e^{-y}dy$$

Then rewrite it as a double integral

$$\Gamma(m)\Gamma(n) = \int_0^\infty \int_0^\infty x^{m-1} y^{n-1} e^{-x-y} dx dy$$

Applying the substitution $x = vt$ and $y = v(1-t)$, we have

$$\Gamma(m)\Gamma(n) = \int_0^1 t^{m-1}(1-t)^{n-1} dt \int_0^\infty v^{m+n-1} e^{-v} dv$$

Using the definitions of Γ and Beta functions, we have

$$\Gamma(m)\Gamma(n) = B(m,n)\Gamma(m+n)$$

Hence proved.

□

A.2 The pdf of Student's t-distribution is a pdf

The values of $t_r(x)$ is positive and the integral is 1.

$$\int_{-\infty}^{\infty} t_r(x) \,\mathrm{d}x = 1$$

Proof. Consider the variable part of Student's t-distribution

$$f(x) = (1 + \frac{x^2}{r})^{-\frac{r+1}{2}}, -\infty < x < \infty$$

use a replacement as following

$$x^2 = \frac{y}{1-y}$$

it is easy to see that $\lim_{y\to 0} x = 0$ and $\lim_{y\to 1} x = \infty$. Additionally, the x^2 is even function. Thus we can write the integral of $f(x)$

$$\int_{-\infty}^{\infty} f(x) \,\mathrm{d}x = 2\sqrt{r} \int_0^1 (\frac{1}{1-y})^{-\frac{r+1}{2}} \,\mathrm{d}(\frac{y}{1-y})^{\frac{1}{2}}$$

it is not hard to find out that the integral may end up with

$$\sqrt{r} \int_0^1 (1-y)^{\frac{r}{2}-1} y^{\frac{1}{2}-1} \,\mathrm{d}y = \sqrt{r} B(\frac{r}{2}, \frac{1}{2})$$

Finally the normalization factor has to be

$$\frac{\Gamma(\frac{r+1}{2})}{\sqrt{r}\Gamma(\frac{r}{2})\Gamma(\frac{1}{2})}$$

which makes the integral of $t_r(x)$ is 1.

□

A.3 Compute the pdf of Student's t-distribution

Here, we provide a simple computation of the pdf of the Student's t-distribution.

$$T = \frac{X}{\sqrt{Y/r}}$$

in which $X \sim N(0,1)$ and $Y \sim \chi^2(r)$, and they are independent. Thus, we have

$$p(x) \propto e^{-x^2/2}$$
$$p(y) \propto y^{r/2-1} \cdot e^{-y/2}$$

The random variable t follows the equation $t = \frac{x}{\sqrt{y/r}}$. Since we want to proof that

$$p(t) \propto (1 + \frac{t^2}{r})^{-\frac{r+1}{2}}$$

Proof. The joint probability of $p(x,y)$ matches

$$p(x,y) \propto e^{-x^2/2} \cdot y^{r/2-1} \cdot e^{-y/2}$$

And the divergence of $p(x,y)$ is $p(x,y)dxdy$. We can use the variable replacement of

$$y = \frac{x^2}{t^2} \cdot r$$
$$\frac{dy}{dt} \propto \frac{x^2}{t^3}$$

Thus we have the joint probability of $p(x,t)$ matches

$$p(x,t) \propto e^{-x^2/2} \cdot (\frac{x^2}{t^2})^{r/2-1} \cdot e^{-\frac{x^2}{2t^2}r} \cdot \frac{x^2}{t^3}$$

The probability of $p(t)$ can be expressed as

$$p(t) \propto \int_x p(x,t)dx$$

Analysis the expression, we have

$$p(t) \propto t^{-r-1} \int_x x^r \cdot e^{-\frac{1}{2}(1+\frac{r}{t^2})x^2} dx$$
$$p(t) \propto t^{-r-1} \cdot (1 + \frac{r}{t^2})^{\frac{-r-1}{2}} \int_z z^r \cdot e^{z^2} dz$$
$$p(t) \propto (t^2 + r)^{-\frac{r+1}{2}}$$
$$p(t) \propto (1 + \frac{t^2}{r})^{-\frac{r+1}{2}}$$

The process uses the integral of Γ function is constant, and r is constant. □

After that, combining with the A.2, we should finally have the pdf function.