2022 SUMMER

업데이터 통계학 스터디

Chapter 5 – Continuous Random Variables 3

Beta Distribution

Beta function

If $\alpha > 0$ and $\beta > 0$,

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Which is called the beta function.

*There's a high probability that it won't reach you right now.

If you study prior/posterior r.v. in Bayes Statistics in the future, the meaning of beta distribution may will be realized.

Beta Distribution

Definition

The continuous random variable X follows a Beta distribution if its PDF is:

$$f_X(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} I_{(0,1)}(x)$$

We denote $X \sim Beta(\alpha, \beta)$

Example 5.1

Suppose that continuous random variable $X \sim Beta(\alpha, \beta)$. Find

1. $E(X^k)$

2. $E(X)$

3. $Var(X)$

Normal Distribution

Gaussian Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}, \quad \int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

Proof)

Normal Distribution

Definition

The continuous random variable X follows a Normal distribution if its PDF is:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}$$

We denote $X \sim N(\mu, \sigma^2)$

NOTE: $E(X) = \mu, Var(X) = \sigma^2$

Normal Distribution

1. All normal curves are bell-shaped with points of inflection at $\mu \pm \sigma$
2. All normal curves are symmetric about the mean μ
3. The area under an entire normal curve is 1
4. All normal curves are positive for all x. That is, $f_X(x) > 0$ for all x.
5. Convergence; $\lim_{x \to \infty} f_X(x) = 0$, and $\lim\limits_{x \to -\infty} f_X(x) = 0$
6. The height of any normal curve is maximized at $x = \mu$

Standard Normal Distribution

Definition

If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

Which is called the Standard Normal Distribution.

Standardization makes it easy to obtain probabilities using a normal distribution table.

Example 5.2

Suppose that continuous random variable $X \sim N(\mu, \sigma^2)$. Find

1. $M_X(t)$
2. $E(X)$
3. $Var(X)$

CDF Method

CDF Method

Let X be a c.r.v.. If we consider $Y = u(X)$ then Y must also be a r.v. that has its own distribution.

We can find the PDF of the random function $Y = u(X)$ by:

1. Find the CDF of Y

$$F_Y(y) = P(Y \leq y)$$

2. Differentiate the CDF of Y

$$f_Y(y) = F'_Y(y)$$

Example 5.3

Let $X \sim N(0,1)$ What is the distribution of $Y = X^2$?

Example 5.4

$X_i \sim Unif(0, \theta)$ independently. Let $Y = max(X_1, \cdots, X_n)$.

What is the distribution of $\frac{Y}{\theta}$?

Change of Variable

Change of Variable

[Continuous, Uni-variate, 1:1 function]

Suppose $Y \sim f_Y(y)$, $U = h(Y)$ where $h(y)$ increasing in y. Then,

$$P(U \leq u) = P(h(Y) \leq u) = P(Y \ \leq h^{-1}(u))$$

Suppose $Y \sim f_Y(y)$, $U = h(Y)$ where $h(y)$ decreasing in y. Then,

$$P(U \leq u) = P(h(Y) \leq u) = P(Y \ \leq h^{-1}(u))$$

$$\therefore f_U(u) = f_Y\left(h^{-1}(u)\right) \left| \frac{dh^{-1}(u)}{du} \right|$$

Example 5.5

Let Y be a c.r.v. with the following PDF:

$$f_Y(y) = 2y\, I_{(0,1)}(y)$$

What is the PDF of $U = 8y^3$?

Example 5.6

Let Y be a c.r.v. with the following PDF:

$$f_Y(y) = 4y^3\, I_{(0,1)}(y)$$

What is the PDF of $U = e^Y$?

MGF Method

MGF Method

As we saw earlier, MGF uniquely determines the distribution.

Let X be a c.r.v.. If we consider $Y = u(X)$ then Y must also be a r.v. that has its own distribution.

We can find the distribution of the random function $Y = u(X)$ by proving that the MGF of Y matches the MGF of a particular distribution .

Example 5.7

Find a distribution;

$X \coloneqq$ Sum of i.i.d. $Exponential(\theta)$

Example 5.8

Find a distribution;

$Y \coloneqq$ Sum of independent $X_i \sim Bin(n_i, p)$

Chi-Square Distribution (Recall)

Chi-Square Distribution

Definition

If $X \sim N(\mu, \sigma^2)$, then:

$$U = (\frac{X-\mu}{\sigma})^2 = Z^2$$

Is distributed as a Chi-Square r.v. with 1 $d.o.f.$ $\Leftrightarrow U \sim \chi^2(1)$

Generalization; $U = {Z_1}^2 + \cdots + {Z_n}^2 \sim \chi^2(n)$ (where Z_i are indep)

Furthermore; $U_1, \cdots, U_n$: independent $\chi^2(r_i) \Rightarrow \sum U_i \sim \chi^2(\sum r_i)$

Chi-Square Distribution

Proof)

t Distribution

t Distribution

Definition

$Z \sim N(0,1)$ and $W \sim \chi^2(v)$, independently. Then

$$T := \frac{Z}{\sqrt{W/v}} \sim t(v)$$

$$f_T(t) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}(\frac{v}{2})}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}} I_{\mathbb{R}}(t)$$

t Distribution

$E(T) = 0, Var(T) = \frac{v}{v-2}\ for\ v > 2$

If v$(d.o.f.)\ increase < \infty \ \Rightarrow \ T \longrightarrow N(0,1)$

F Distribution

F Distribution

Definition

$W_1 \sim \chi^2(v_1)$ and $W_2 \sim \chi^2(v_2)$, independently. Then

$$\mathrm{F} \coloneqq \frac{W_1/V_1}{W_2/V_2} \sim F(v_1, v_2)$$

$$PDF?\ You\ don't\ have\ to\ know.$$

$$f(y)=\frac{\Gamma((\nu_1+\nu_2)/2)(\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}y^{(\nu_1/2)-1}\left(1+\frac{\nu_1 y}{\nu_2}\right)^{-(\nu_1+\nu_2)/2}, \quad 0<y<\infty$$

F Distribution

$E(F) = \frac{v_2}{v_2 - 2}$ for $v_2 > 2$

$$Var(F) = \frac{2{v_2}^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)} \; for \;\; v_2 > 4$$

- $F \sim F(v_1, v_2) \Longrightarrow \frac{1}{F} \sim F(v_2, v_1)$