2022 SUMMER

업데이터 통계학 스터디

Chapter 4 – Continuous Random Variables 2

Uniform Distribution

Uniform distribution

Definition

A continuous random variable X has a uniform distribution, denoted $X \sim U(a, b)$, if its probability density function is:

$$f_X(x) = \frac{1}{b-a}, \qquad (a < x < b)$$

For $a, b \in \mathbb{R},\ a \neq b$.

Restricting $a = 0$ and $b = 1$, the resulting distribution $U(0,1)$ is called a standard uniform distribution.

Uniform distribution

The cumulative distribution function of a uniform random variable X is:

$$F_X(x) = \frac{x-a}{b-a}, \qquad (a < x < b)$$

For $a, b \in \mathbb{R},\ a \neq b$.

As the picture shows $F_X(x) = 0$ when $x < a$ and $F_X(x) = 1$ when $x > b$. The slope of the line between a and b is, of course, $1/(b-a)$.

Uniform distribution

For a continuous uniform random variable X defined over the support $a < x < b$, that is, $X \sim U(a, b)$:

- $\mu_X = E(X) = \frac{a+b}{2}$
- ${\sigma^2}_X = Var(X) = \frac{(b-a)^2}{12}$
- $M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$

Example 4.1

Students arrive randomly at a Up-data statistics study. Given that one student arrived during a particular 10 minutes period, let X be the time within the 10 mins that the student arrived. If $X \sim U(0, 10)$, find

1. PDF of X
2. $P(X > 8)$
3. $P(2 < X < 8)$
4. $E(X)$
5. $Var(X)$
6. $M_X(t)$

Poisson Process

Poisson Process

We have learned that Poisson random variable represents several rare events. Let's derive the formula under the following condition.

Let N denote the number of events occurred in a given continuous interval. Then N follows an approximate Poisson process with parameter $\lambda > 0$ if:

Poisson Process

- The number of events occurring in non-overlapping sub-intervals are independent.
- The probability of exactly one event in a short subinterval of length $h = 1/n$ is approximately $\lambda h = \lambda(1/n) = \lambda/n$.

  - If the length of the interval is small, the probability of the event is also small.

- The probability of exactly two or more events in a short subinterval is essentially zero.

  - It says that the event rarely occurs.

Poisson Process

Let $N(t)$ is the number of rare events occurred over the period $[0, t]$.

For example, let X be the number of cars passing through an intersection within 1 minute. Assume that the rate is λ per 1 minute. Then $X = N(1)$.

Goal: Find $P(X = x)$ for a positive integer x.

Poisson Process

Let $N(t)$ be a Poisson process with a rate λ.

Then, $N(t) \sim Poisson(\lambda t)$

Exponential Distribution

Exponential Distribution

Let $N(t)$ be the number of customers coming at a bank in an interval of length t. Assuming it following Poisson process with a rate λ per each interval of length 1. Then, $N(t) \sim Poisson(\lambda t)$.

Let X be the waiting time until the first customer arrives at the bank. Find its CDF, PDF, and the expected waiting time.

Exponential Distribution

Because the waiting time is nonnegative, $F_X(t) = 0$ when $t \leq 0$, where t is a waiting time. Given $t > 0$,

$$
\begin{aligned}
F_X(t) &= P(\mathit{waiting\ time\ is\ less\ than\ t}) \\
&= P(X \leq t) \\
&= 1 - P(X > t) \\
&= 1 - P(\mathit{there\ is\ no\ event\ until\ t}) \quad \cdots \quad (1)
\end{aligned}
$$

Exponential Distribution

By $N(t) \sim Poisson(\lambda t)$, Equation (1) becomes like below.

$$\begin{aligned} F_X(t) &= 1 - P(there\ is\ no\ event\ until\ t) \\ &= 1 - P(N(t) = 0) \\ &= 1 - e^{-\lambda t} \end{aligned}$$

Therefore, the waiting time for the first customer has CDF $F_X(x) = 1 - e^{-\lambda x}$ $(x > 0)$. Then, $X \sim Exp(\lambda)$ and its PDF is

$$f_X(x) = \lambda e^{-\lambda x}\, I_{(0,\infty)}(x)$$

Exponential Distribution

The continuous random variable X follows an exponential distribution if its PDF is:

$$f_X(x) = \lambda e^{-\lambda x}\, I_{(0,\infty)}(x), \quad \text{for } \lambda > 0$$

We denote $X \sim Exp(\lambda)$

Example 4.2

Suppose that continuous random variable $X \sim Exp(\lambda)$. Find

1. $M_X(t)$

2. $E(X)$

3. $Var(X)$

4. $F_X(x)$

5. $P(X > x)$

Example 4.3

Students arrive at a local bar according to a Poisson process at a mean rate of 30 students per hour. What is the probability that the bouncer must wait more than 3 minutes to card the next student?

Exponential Distribution

Memoryless property

Let $X \sim Exp(\lambda)$. Compute

$$P(X \geq t + t_0 \mid X \geq t_0)$$

Suppose that X is a waiting time until the first customer.

A person has waited for t_0 minutes so far. What is the probability that the person will wait for $t + t_0$ minutes?

Example 4.4

The number of miles that a particular car can run before its battery wears out is exponentially distributed with an average of 10,000 miles.

The owner needs to take a 5,000 miles trip. What is the probability that he will be able to complete the trip without replacing the battery?

Gamma Distribution

Gamma Function

Definition

The gamma function, denoted $\Gamma(t)$, is defined, for $t > 0$, by:

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy$$

1. $\Gamma(t) = (t-1)\Gamma(t-1),\ for\ \ t > 1$
2. $\Gamma(n) = (n-1)!\,,\ \ if\ n \in \mathbb{N}$
3. $\Gamma(1) = 1,\ \ \ \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Gamma Distribution

In a Poisson process with mean λ, the waiting time X until the first event occurs follows an exponential distribution with mean $1/\lambda$.

Let X_α denote the waiting time until the αth event occurs.

Find the distribution of X_α.

Gamma Distribution

The CDF of X_α when $x \geq 0$ is given by

$$
\begin{aligned}
F_X(x) &= P(X_\alpha \leq x) \\
&= 1 - P(X_\alpha > x) \\
&= 1 - P(fewer\ than\ \alpha\ occurrences\ in\ [0, x]) \\
&= 1 - \sum_{k=0}^{\alpha-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}
\end{aligned}
$$

Gamma Distribution

since the number of occurrences in the interval [0,x] has a Poisson distribution with mean λx.

Then, PDF of X_α:

$$f_X(x) = F'_X(x) = \frac{\lambda e^{-\lambda x} \lambda x^{\alpha-1}}{(\alpha-1)!}, \quad x > 0$$

Gamma Distribution

Recall: $\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy$, $\Gamma(n) = (n-1)!$

$$f_X(x) = \frac{\lambda e^{-\lambda x} \lambda x^{\alpha-1}}{(\alpha-1)!} = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad for\ x > 0$$

where $\beta = \frac{1}{\lambda}$. Then,

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

Gamma Distribution

A continuous random variable X follows a gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if its PDF is:

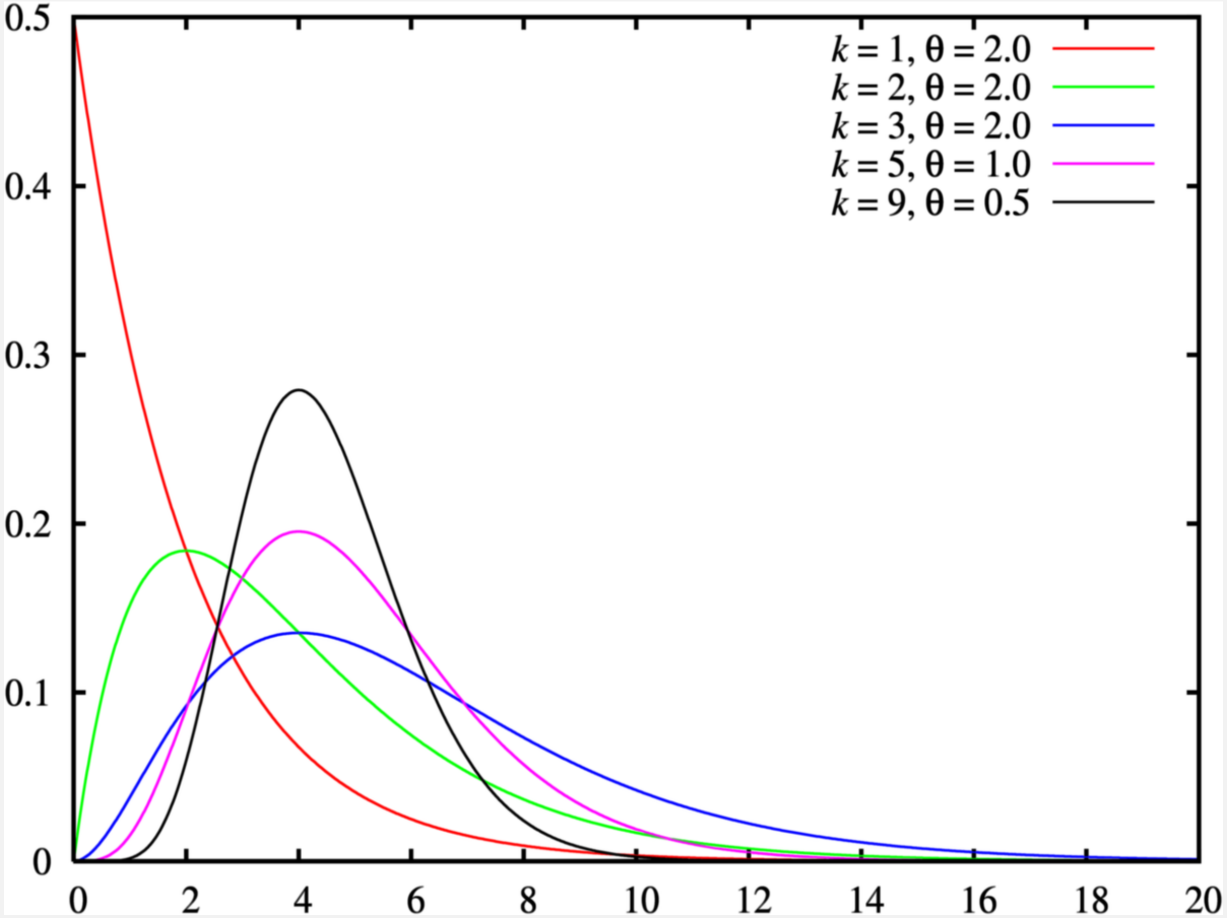
$$f_X(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} \ I_{(0,\infty)}(x)$$

Then we denote $X \sim Gamma(\alpha, \beta)$

- $\alpha(or\ k)$: shape parameter, $\beta(or\ \theta)$: scale parameter,

  $\lambda = \frac{1}{\beta}$: rate parameter

Gamma Distribution



Example 4.5

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

1. $M_X(t)$

2. $E(X)$

3. $Var(X)$

Example 4.6

Telephone calls arrive at Jiwon's phone at a mean rate of λ = 2 per-minute according to a Poisson process. Let X denote the waiting time in minutes until the fifth call arrives.

1. $f_X(x)$
2. $E(X)$
3. $Var(X)$

Example 4.6

Suppose the number of customers per hour arriving at a shop follows a Poisson process with mean 30. That is, if a minute is our unit, $\lambda = 1/2$.

What is the probability that the shopkeeper will wait more than 5 minutes before both first two customers arrive?

Chi-Square Distribution

Chi-Square Distribution

Definition

X follow a gamma distribution with $\beta = 2\ and\ \alpha = \frac{v}{2}, v \in \mathbb{N}$.

Then the PDF of X is:

$$f_X(x) = \frac{1}{\Gamma(v/2)2^{v/2}} x^{(v/2)-1} e^{-\frac{x}{2}}\ I_{(0,\infty)}(x)$$

Then we denote $X \sim Gamma\left(\frac{v}{2}, 2\right) \coloneqq \chi^2(v)$, v is a $d.o.f$.

* Later, the Chi-Square distribution is related to normal r.v.

Example 4.7

Suppose that continuous random variable $X \sim \chi^2(v)$. Find

1. $M_X(t)$

2. $E(X)$

3. $Var(X)$

Chi-Square Distribution

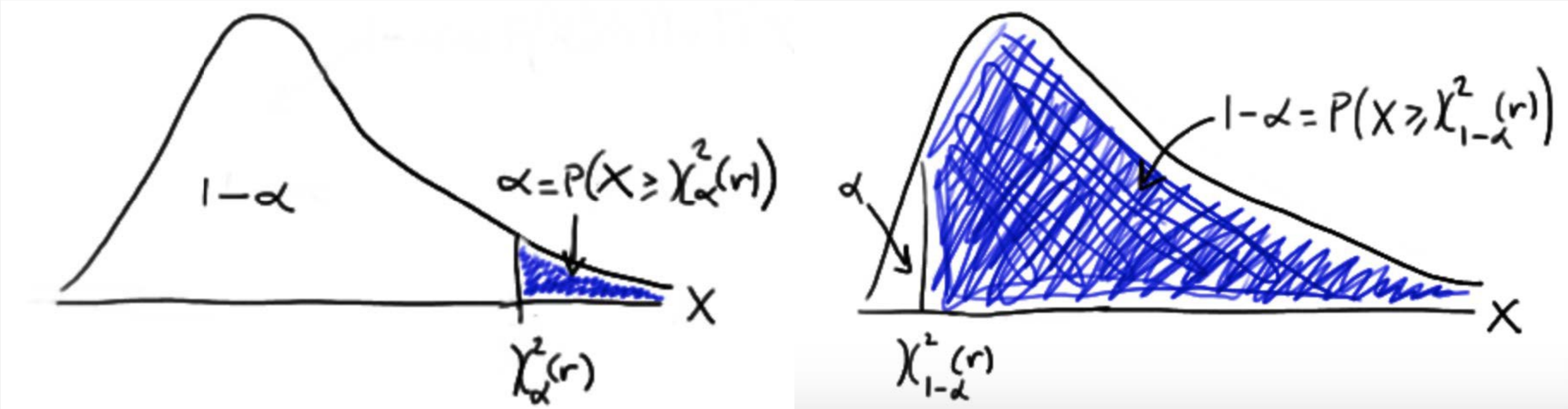
Table

Let α be a positive probability between 0 and 1 and let X have a chi-square distribution with v degrees of freedom. Then,

The upper $100\alpha^{th}$ percentile is the value ${\chi_\alpha}^2(v)$ such that the area under the curve and to the right of ${\chi_\alpha}^2(v)$ is α.

That is, $P[X \geq {\chi_\alpha}^2(v)] = \alpha \Leftrightarrow P[X \leq {\chi_{1-\alpha}}^2(v)] = \alpha$

Chi-Square Distribution



Example 4.8

If customers arrive at a shop on the average of 30 per hour in accordance with a Poisson process.

what is the probability that the shopkeeper will have to wait longer than 9.390 minutes for the first nine customers to arrive?

Example 4.9

If X has an exponential distribution with a mean of 2.

Find $P(0.051 < X < 7.378)$