VE401 Recitation Class Note3

Continuous Random Variables

Chen Siyi
siyi.chen_chicy@sjtu.edu.cn

1 Overview

1.1 Definiation

Let S be a sample space. A continuous random variable is a map

$$X : S \to \mathbb{R}$$

together with a probability **density** function

$$f_X : \mathbb{R} \to \mathbb{R}$$

where

(i) $f_X(x) \geq 0$ for all $x \in \mathbb{R}$

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

1.2 General Properties

Various distributions' important features:

1. f_X: probability density function(PDF). Notice in comparison to the discrete random variables, now $f_X \neq P[X = x]$. $P[X = x] = 0$.
2. $F_X(x) = \int_{-\infty}^{x} f_X(y)\,\mathrm{d}y$: cumulative distribution function(CDF). And $F'_X(x) = f_X(x)$ holds.
3. $\mathrm{E}[X] := \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x$: expectation.
4. $\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = E[X^2] - E[X]^2$: Variance.
5. $m_X(t) := E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,\mathrm{d}x$: moment generating function (MGF).

2 Locations

1. The median M_x: $P[X \leqslant M_x] = 0.5$
2. The mean $E[X]$: The average value.
3. The mode x_0: The location having the maximum f_X (if there is a unique maximum location).

3 Memoryless

Definiation:

$$P[X > x + s|X > x] = P[X > s]$$

Interpolation:

Let's observe the definition above. What do you notice? P[X>x+s | X>x]=P[X>s] does not rely on "x". We can interpret as: P[X>x+s | X>x] does not "remember" that it is already with X>x.

A concrete example:

There is a type of machine. The probability of the machine's service life follows an exponential distribution, which we know is memoryless.

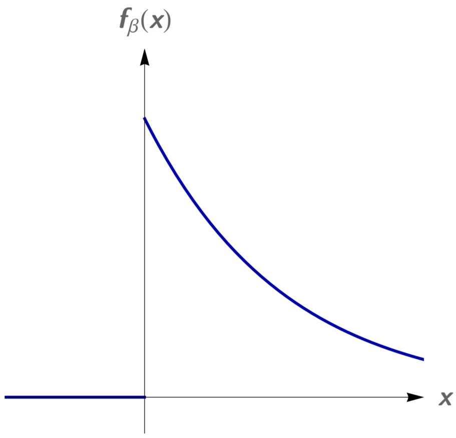


Figure 1: The exponential distribution

Then P[X>x+s | X>x]=P[X>s] simply means: Given the machine being used for x time, then the probability for it being able to work for another s time, is just the same as a new machine being able to work for s time. This means, the machine does not remember that it has been used for x time already.

Comments:

Memoryless is a property not only held by some continuous random variables like the exponential distribution, but also shared by some discrete random variables.

Question1: Memoryless

Determine whether the following discrete random variables are memoryless:

(i) Poisson distribution

(ii) Geometric distribution

(iii) Binomial distribution

How do you interpret the above results?

4 Specific Distributions

4.1 Exponential Distribution

Connection to the Poisson Distribution:

Start from the point of an arrival, the time for one successive arrival of a Poisson-distributed random variable to occur is exponentially distributed with parameter $\beta = \lambda$. (Recall: $k = \lambda t$)

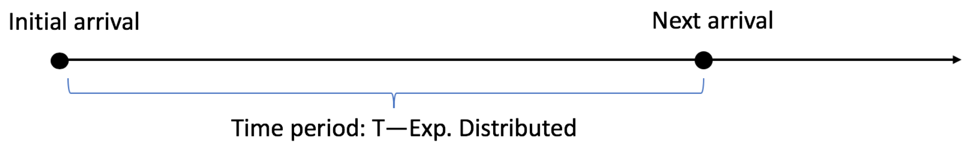


Figure 2: The exponential distribution

Features:

1. β is the parameter.
2. $f_\beta(x) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$
3. $F_X(x) = 1 - e^{-\beta t}$
4. E[X]=$\frac{1}{\beta}$
5. Var[X]=$\frac{1}{\beta^2}$
6. $m_X(t) = (1 - \frac{t}{\beta})^{-1}$
7. Memoryless

Question2: "Contradictory" Memoryless Problem

You arrive at the bus stop at 10 o'clock, knowing that bus arrivals follow a Poisson distribution with a rate of 2 buses per hour.

(i) What is the probability that you will have to wait longer than 10 minutes?

(ii) If, at 10:15, the bus has not yet arrived, what is the probability that you will have to wait at least an additional 10 minutes?

Solving this problem(very simple, so it's left to you, and you are also encouraged to), you get the exact same answer(around 0.717) for both cases. Notice we say the Poisson distribution is not memoryless. So it causes "contradiction" now. Why?

4.2 Gamma Distribution

Connection to the Poisson Distribution:

Start from the point of an arrival, the time for r successive arrivals of a Poisson-distributed random variable to occur follows a gamma distribution with parameter $\alpha = r$, $\beta = \lambda$. (Recall: $k = \lambda t$)

So it's also a generalization of the exponential distribution.

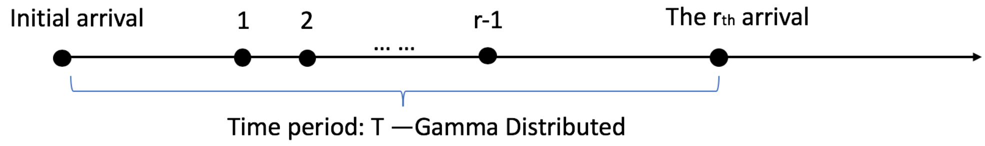
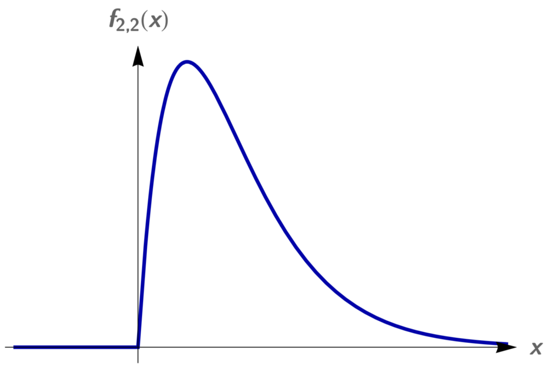


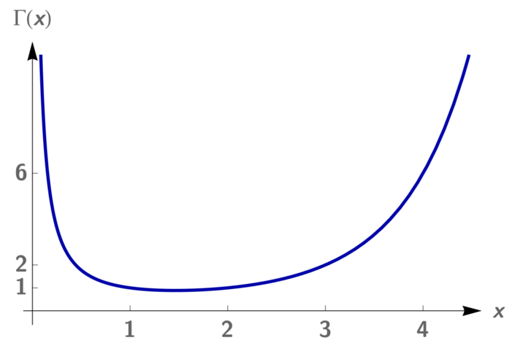
Figure 3: The gamma distribution

Features:

1. α, β are the parameters.
2. $f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$
3. $F_X(x) = 1 - e^{-\beta t}$
4. E[X]=$\frac{\alpha}{\beta}$
5. Var[X]=$\frac{\alpha}{\beta^2}$
6. $m_X(t) = (1 - \frac{t}{\beta})^{-\alpha}, \quad m_X : (-\infty, \beta) \to \mathbb{R}$
7. Not memoryless



(a) The gamma distribution



(b) The gamma function

Figure 4: Gamma

Euler Gamma Function:

1. $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1}e^{-z}dz$
2. $\Gamma(1) = 1$, $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ for $\alpha > 1$
3. $n! = \Gamma(n+1)$, for $n \in \mathbb{R}$

Question3: Euler Gamma Function

Find: $\Gamma(\frac{1}{2})$, $\Gamma(\frac{2n+1}{2})$

4.3 Chi-squared Distribution

Connection to the Gamma Distribution:
It's a special case of the gamma distribution, with $\alpha = \frac{\gamma}{2}$, $\beta = 2$, where $\gamma \in \mathbb{N}$

Comments:
We say (χ^2_γ, f_X) follows a chi-squared distribution with γ degree of freedom. In this special distribution, X is written as χ^2_γ. See the reasons and other important points later.
A very important distribution.

Features:

1. γ is the parameter.
2. $f_\gamma(x) = \begin{cases} \frac{1}{\Gamma(\gamma/2)2^\alpha}x^{\gamma/2-1}e^{-x/2}, & x > 0 \\ 0, & x \le 0 \end{cases}$
3. $F_X(x) = 1 - e^{-\beta t}$
4. E[χ^2_γ]=γ
5. Var[χ^2_γ]=2γ
6. $m_X(t) = (1-\frac{t}{2})^{-\frac{\gamma}{2}}, \quad m_X : (-\infty, 2) \to \mathbb{R}$

4.4 Normal(Gauß) Distribution

Definition:

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

Features:

1. μ, δ are the parameters
2. E[X] $= \mu$
3. Var[X] $= \delta^2$

4. $m_X(t) = e^{\mu t + \sigma^2 t^2/2}$

5. $$P[-\sigma < X - \mu < \sigma] = 0.68$$
$$P[-2\sigma < X - \mu < 2\sigma] = 0.95$$
$$P[-3\sigma < X - \mu < 3\sigma] = 0.997$$

4.5 Standard Normal Distribution

Definition:

Let X be a normally distributed random variable with mean μ and standard deviation δ. Then $Z = \frac{X-\mu}{\delta}$ follows a standard normal distribution with mean 0 and variance 1.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Features:

1. E[Z] = 0
2. Var[Z] = 1
3. $m_Z(t) = e^{t^2/2}$

CDF:

$$\Phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt = \frac{1}{2} erfc(-\frac{z}{\sqrt{2}})$$

Where we define:

$$\mathrm{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt, \quad \mathrm{erfc}(z) := 1 - \mathrm{erf}(z)$$

Approximate the Binomial:

$$P[X \leq y] = \sum_{x=0}^{y} \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x} \approx \Phi\left(\frac{y + 1/2 - np}{\sqrt{np(1-p)}} \right)$$

<u>Be careful with the half-unit correction.</u>

Relation to the Chi-square:

A Chi-square distributed variable with γ degree of freedom, is the sum of r square of standard normal distributed variables. Simply:

$$\chi_n^2 = \sum_{i=1}^{n} Z_i^2$$

5 Transformation of Random Variables

Let X be a continuous random variable with density f_X.
Let Y $= \varphi \circ X$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly monotonic and differentiable. The density for Y is then given by

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|\tfrac{d\varphi^{-1}(y)}{dy}\right| \qquad \text{for } y \in \operatorname{ran}\varphi$$
$$f_Y(y) = 0 \qquad \text{for } y \notin \operatorname{ran}\varphi$$

It is important to know how to prove.

Step 1: Find CDF $F_Y(y)$.

$$\begin{aligned} F_Y(y) &= P[\varphi(X) \leq y] \\ &= P\left[\varphi^{-1}(\varphi(X)) \geq \varphi^{-1}(y)\right] \\ &= P\left[X \geq \varphi^{-1}(y)\right] \\ &= 1 - P\left[X \leq \varphi^{-1}(y)\right] \\ &= 1 - F_X\left(\varphi^{-1}(y)\right) \end{aligned}$$

Step 2: Find $f_Y(y) = F'_Y(y)$

$$\begin{aligned} f_Y(y) = F'_Y(y) &= -f_X\left(\varphi^{-1}(y)\right) \frac{d\varphi^{-1}(y)}{dy} \\ &= f_X\left(\varphi^{-1}(y)\right) \cdot \left|\frac{d\varphi^{-1}(y)}{dy}\right| \end{aligned}$$

Question: Transform Variables

Assume X is a continuous random variable with PDF f_X, $-\infty < x < \infty$.
Find PDF for the random variable Y, where $Y = X^2$.
Find PDF for Y when X $\sim$ N(0,1), which is a standard normal distribution.

Answer: Transform Variables

When $y \leqslant 0$, $F_Y(y) = 0$. When $y > 0$:

$$\begin{aligned} F_Y(y) &= P\{Y \leqslant y\} = P\left\{X^2 \leqslant y\right\} \\ &= P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}\} \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

Differentiate:

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right], & y > 0 \\ 0, & y \leqslant 0 \end{cases}$$

And for X $\sim$ N(0,1):

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} y^{-1/2} \mathrm{e}^{-y/2}, & y > 0 \\ 0, & y \leqslant 0 \end{cases}$$

This is exactly a Chi-square distribution with $\gamma = 1$.

6 The Chebyshev Inequality

6.1 Definition

Let c $>$ 0 be any real number, and for k $\in \mathbb{N} - 0$, then for any random variables:

$$P[|X| \geq c] \leq \frac{E\left[|X|^k\right]}{c^k}$$

6.2 Application

In general, for any variables:

$$\begin{gathered} P[|X - \mu| \geq m\sigma] \leq \frac{1}{m^2} \\ P[-m\sigma < X - \mu < m\sigma] \geq 1 - \frac{1}{m^2} \end{gathered}$$

7 Answers

Answer1: Memoryless

Determine whether the following discrete random variables are memoryless:

(i) Poisson distribution: Not memoryless.

(ii) Geometric distribution: Memoryless. And it is the only memoryless discrete random variable.

(iii) Binomial distribution: Not memoryless.

Interpret:

(i) Poisson distribution: Within a fixed interval, the more you gain before, the less you are expected to gain after.

(ii) Geometric distribution: All the trials are independent. The previous results won't affect the following ones. Recall flowing a coin.

(iii) Binomial distribution: Similar to the Poisson Distribution.

Answer2: "Contradictory" Memoryless Problem

The bus arrivals follow the Poisson distribution which is not memoryless, and the waiting time follows the Exponential distribution which is memoryless.
The problem is asking about the waiting time, so there's actually no contradiction.
By the way, how will you design a problem asking for the Poisson distribution based on this background?

Answer3: Euler Gamma Function

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$,
$\Gamma(\frac{2n+1}{2}) = \frac{(2n-1)(2n-3)...1}{2^n}\sqrt{\pi}$