7. 常用連續機率分佈

Common Continuous Probability Distributions

- 連續均勻分佈。
- 常態分佈。
- 伽瑪分佈。
- 指數分佈。

7.1 連續均勻分佈

連續均勻分佈 (Continuous Uniform Distribution)

連續均勻隨機變數 X 在區間 $[A, B]$ 上的密度函數定義為 $f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A \le x \le B \\ 0, & \text{elsewhere} \end{cases}$

What does a continuous uniform function look like?

範例7.1

- Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval [0, 4]. (a) What is the probability density function? (b) What is the probability that any given conference lasts at least 3 hours?

定理 7.1

How to prove it?

在定義區間 $[A, B]$ 之連續均勻隨機變數 X 的平均值 $\mu = \frac{A+B}{2}$，變異數 $\sigma^2 = \frac{(A-B)^2}{12}$。

7.2 常態分佈

常態分佈 (Normal Distribution)

常態隨機變數 X 之密度函數定義為 $n(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-1}{2\sigma^2}(x-\mu)^2}, -\infty < x < \infty$

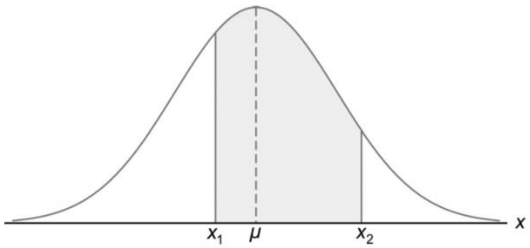
What does a normal distribution function look like?

定理 7.2

How to prove it?

常態分佈 $n(x;\mu,\sigma)$ 的平均值為 μ 且變異數為 σ^2，其標準差為 σ。

7.3 常態分佈的應用



- What does the area of the shaded region represent?
- How to calculate such area? Any easier way?
- How many different normal distributions are there?
- Can we use "only one normal distribution" to obtain such area under "different normal distribution curves"?

定義 7.1
若常態隨機變數的平均值為0且標準差為1，我們稱之為標準常態隨機變數，且其分佈稱為標準常態分佈 (standard normal distribution)。

- How to convert a "normal random variable" to a "standard normal random variable"?

範例7.2

- Given a standard normal distribution Z, find the area under the curve that lies (a) to the right of z = 1.84 and (b) between z = −1.97 and z = 0.86.

範例7.3

- Given a standard normal distribution Z, find the value of k such that (a) $P(Z > k) = 0.3015$ and (b) $P(k < Z < -0.18) = 0.4197$.

範例7.4

- Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$, find the probability that X assumes a value between 45 and 62.

範例7.5

- Given a random variable X having a normal distribution, find $P(\mu-2\sigma < X < \mu+2\sigma)$. Compare it with the result obtained using Chebyshev's theorem.

範例7.6

- Given a normal distribution with $\mu = 40$ and $\sigma = 6$, find the value of x that has (a) 45% of the area to the left and (b) 14% of the area to the right.

範例7.7

- In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01 cm. It is known that in this process, the diameter of a ball bearing has a normal distribution with mean $\mu = 3.0$ and standard deviation $\sigma = 0.005$. On average, how many manufactured ball bearings will be scrapped?

範例7.8

- Gauges are used to reject all components for which a certain dimension is not within the specification $1.50 \pm d$. It is known that this measurement is normally distributed with mean 1.50 and standard deviation 0.2. Determine the value d such that the specifications cover 95% of the measurements.

範例7.9

- The average grade for an exam is 74, and the standard deviation is 7. If 12% of the class is given As, and the grades are curved to follow a normal distribution, what is the lowest possible A and the highest possible B?

範例7.10

- Refer to Example 7.9 and find the sixth decile.

7.4 近似二項式分佈的常態分佈

定理 7.3

假設隨機變數 X 為一平均值為 $\mu = np$、變異數為 $\sigma^2 = npq$ 之二項式分佈。考慮以下隨機變數 Z

$$Z = \frac{X - np}{\sqrt{npq}}$$

What distribution does Z originally have?

當 $n \to \infty$ 時，Z會趨近於標準常態分佈 $n(z\,;0\,,1)$

7.4 近似二項式分佈的常態分佈(contd.)

- Let X be a binomial random variable with parameters n and p. For large n, X has approximately a normal distribution with $\mu = np$ and $\sigma^2 = npq = np(1-p)$ and

$$\begin{aligned} P(X \le x) &= \sum_{k=0}^{x} b(k; n, p) \\ &\approx \text{area under normal curve to the left of } x + 0.5 \\ &= P\left(Z \le \frac{x + 0.5 - np}{\sqrt{npq}}\right), \end{aligned}$$

the approximation will be good if np and n(1-p) are greater than or equal to 5

範例7.11

- A multiple-choice quiz has 200 questions, each with 4 possible answers of which only 1 is correct. What is the probability that sheer guesswork yields from 25 to 30 correct answers for the 80 of the 200 problems about which the student has no knowledge?

7.5 伽瑪分佈及指數分佈

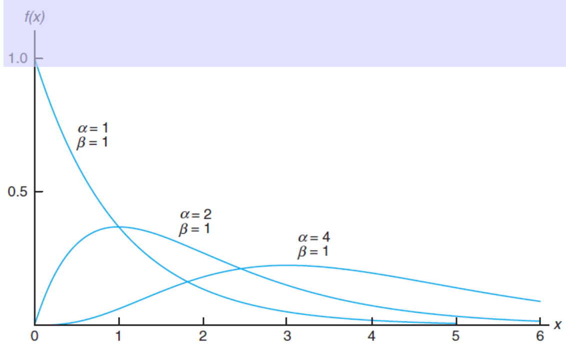
定義 7.2

伽瑪函數(Gamma function)定義：$\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}e^{-x}dx$, for $\alpha>0$

- 伽瑪函數的一些特性：
  - 對於正整數 n 而言，$\Gamma(n)=(n-1)(n-2)\cdots(1)\Gamma(1)$
  - 對於正整數 n 而言，$\Gamma(n)=(n-1)!$
  - $\Gamma(1)=1$ 。
  - $\Gamma(1/2)=\sqrt{\pi}$ 。

伽瑪分佈 (Gamma Distribution)

若連續隨機變數 X 是參數 $\alpha>0$ 和 $\beta>0$ 的伽瑪分佈，則其密度函數為 $f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, & x>0\\ 0, & \text{elsewhere}\end{cases}$



- The special gamma distribution for which α = 1 is called the exponential distribution.

7.5 伽瑪分佈及指數分佈(contd.)

指數分佈 (Exponential Distribution)

若連續隨機變數 X 是參數 $\beta > 0$ 的指數分佈，則其密度函數為 $f(x;\beta) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0 & \text{, elsewhere} \end{cases}$

定理 7.4

伽瑪分佈的平均值 $\mu = \alpha\beta$，變異數 $\sigma^2 = \alpha\beta^2$。

推論 6.1

指數分佈的平均值 $\mu = \beta$，變異數 $\sigma^2 = \beta^2$。

- Relationship to the Poisson Process
  - Let X be the time to the first Poisson event.
  - X has an exponential distribution with $\lambda = 1/\beta$ Why?

範例7.12

- Suppose that a system contains a certain type of component whose time, in years, to failure is given by T. The random variable T is modeled nicely by the exponential distribution with mean time to failure β = 5. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

- The memoryless property of the exponential distribution
  - For example, an electronic component where lifetime has an exponential distribution, the probability that the component lasts t hours is the same as the conditional probability

    $P(X \geq t_0 + t \mid X \geq t_0) = P(X \geq t)$ What does the equation mean? How to derive it?

- The exponential distribution describes the time elapsed until the occurrence of a Poisson event, or the time between Poisson events.

- The gamma distribution describes the time elapsed until a specified number of Poisson events occur.

範例7.13

- Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse by the time 2 calls have come in to the switchboard?

範例7.14

- In a biomedical study with rats, a dose-response investigation is used to determine the effect of the dose of a toxicant on their survival time. The toxicant is one that is frequently discharged into the atmosphere from jet fuel. For a certain dose of the toxicant, the study determines that the survival time, in weeks, has a gamma distribution with $\alpha = 5$ and $\beta = 10$. What is the probability that a rat survives no longer than 60 weeks?

範例7.15

- It is known, from previous data, that the length of time in months between customer complaints about a certain product is a gamma distribution with $\alpha = 2$ and $\beta = 4$. Changes were made to tighten quality control requirements. Following these changes, 20 months passed before the first complaint. Does it appear as if the quality control tightening was effective?

範例7.16

- The time Y in years before a major repair is required for a certain washing machine is characterized by the density function shown below. The machine is considered a bargain if it is unlikely to require a major repair before the sixth year. What is the probability $P(Y > 6)$? Can this washing machine be considered a bargain? What is the probability that a major repair is required in the first year?

$$f(y) = \begin{cases} \frac{1}{4} e^{-y/4}, & y \ge 0 \\ 0, & \text{elsewhere} \end{cases}$$