Chapter 5 part 1

Continuous Random Variables

Jingchen (Monika) Hu

Vassar College

MATH 241

Outline

Outline

Continuous random variables

- ☞ A *continuous random variable* X can take any real value in $(-\infty, \infty)$.

Continuous random variables

☞ A *continuous random variable* X can take any real value in $(-\infty, \infty)$.

Definition

X is a continuous random variable if there exists a **nonnegative** *function f defined for any $x \in (-\infty, \infty)$, such that for any set B of real numbers,*

$$P(X \in B) = \int_B f(x)dx$$

This function f is called the probability density function (pdf) of the random variable X.

Continuous random variables

☞ A *continuous random variable* X can take any real value in $(-\infty, \infty)$.

Definition

X is a continuous random variable if there exists a **nonnegative** *function f defined for any $x \in (-\infty, \infty)$, such that for any set B of real numbers,*

$$P(X \in B) = \int_B f(x)dx$$

This function f is called the probability density function (pdf) of the random variable X.

Examples of continuous random variable

- Rainfall amount for a year.
- Lifetime of your first car.
- Amount of beer consumed on a game day.

Pdf and cdf

- For pdf to be valid, in additional to being non-negative,

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty < X < \infty) = 1$$

Pdf and cdf

- For pdf to be valid, in additional to being non-negative,

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty < X < \infty) = 1$$

- Cdf function of continuous random variable

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

Pdf and cdf

- For pdf to be valid, in additional to being non-negative,

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty < X < \infty) = 1$$

- Cdf function of continuous random variable

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

- For continuous random variable, probability to be a single point is zero.

$$P(X = a) = \int_{a}^{a} f(x)dx = 0$$

$$P(X < a) = P(X \leq a) - P(X = a) = F(a)$$

Pdf and cdf

- For pdf to be valid, in additional to being non-negative,

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty < X < \infty) = 1$$

- Cdf function of continuous random variable

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

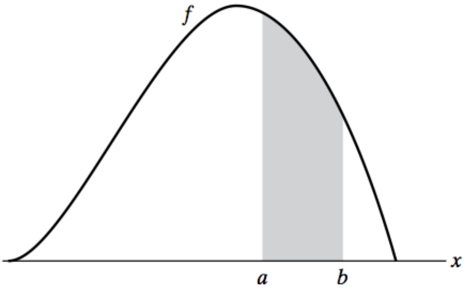
- For continuous random variable, probability to be a single point is zero.

$$P(X = a) = \int_{a}^{a} f(x)dx = 0$$

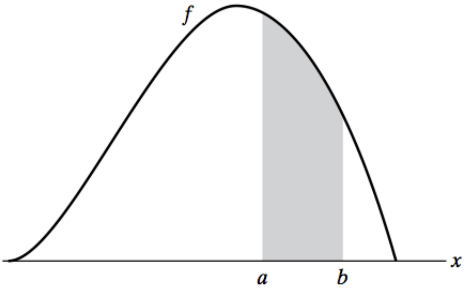
$$P(X < a) = P(X \leq a) - P(X = a) = F(a)$$

- Probability on an interval

$$P(a \leq X \leq b) = F(b) - F(a) = P(a < X < b)$$



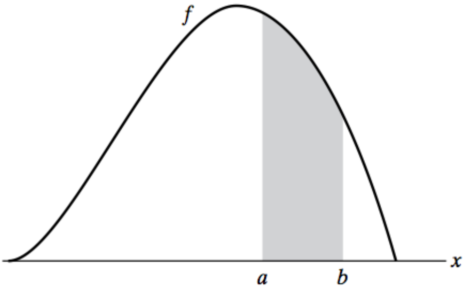
$P(a \leq X \leq b) =$ area of shaded region



$P(a \le X \le b)$ = area of shaded region

Connection between pdf and cdf of continuous random variable. If we know the pdf,

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

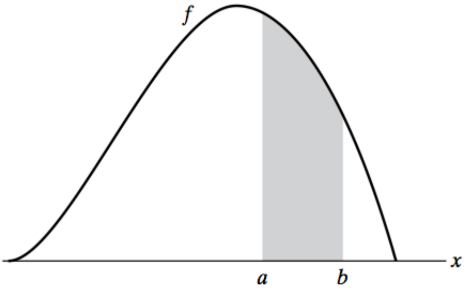


$P(a \leq X \leq b)$ = area of shaded region

Connection between pdf and cdf of continuous random variable. If we know the pdf,

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

How to find the pdf if we know the cdf?



$P(a \leq X \leq b)$ = area of shaded region

Connection between pdf and cdf of continuous random variable. If we know the pdf,

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

How to find the pdf if we know the cdf?

$$f(x) = \frac{d}{dx}F(x)$$

Example: suppose that X is a continuous random variable whose pdf is

$$f(x) = \begin{cases} c(8x - 4x^3) & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- What's the value of c?

Example: suppose that X is a continuous random variable whose pdf is

$$f(x) = \begin{cases} c(8x - 4x^3) & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- What's the value of c?

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x)dx \\ &= \int_0^1 c(8x - 4x^3)dx \\ &= c\left(4x^2 - x^4\Big|_0^1\right) \end{aligned}$$

$$3c = 1 \Longrightarrow c = \frac{1}{3}$$

Interpretation of the pdf

For some small value $h > 0$,

$$
\begin{aligned}
P\left(x - \frac{h}{2} \leq X \leq x + \frac{h}{2}\right) &= \int_{x-\frac{h}{2}}^{x+\frac{h}{2}} f(t)dt \\
&\approx \left[(x + \frac{h}{2}) - (x - \frac{h}{2})\right] f(x) \\
&= h \cdot f(x)
\end{aligned}
$$

The larger $f(x)$ is, the more likely X is to be “near” x.

Recap

A continuous random variable X can take more than countable number of values in $\mathbb{R}$.

- We defined continuous random variable using pdf

$$P(X \in B) = \int_B f(x)dx$$

Recap

A continuous random variable X can take more than countable number of values in $\mathbb{R}$.

- We defined continuous random variable using pdf

$$P(X \in B) = \int_B f(x)dx$$

- Cdf function of continuous random variable

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

$$f(x) = \frac{d}{dx}F(x)$$

Outline

Expectation and variance of continuous random variable

- Recall that expected value of the discrete random variable X with pmf $p(x)$,

$$E[X] = \sum_{\text{all } x} xp(x)$$

Expectation and variance of continuous random variable

- Recall that expected value of the discrete random variable X with pmf $p(x)$,

$$E[X] = \sum_{\text{all } x} x p(x)$$

- We define expected value of a continuous random variable X with pdf $f(x)$ as

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Expectation and variance of continuous random variable

- Recall that expected value of the discrete random variable X with pmf $p(x)$,

$$E[X] = \sum_{\text{all } x} xp(x)$$

- We define expected value of a continuous random variable X with pdf $f(x)$ as

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

- Definitions of variance and standard deviation are the same.

$$Var(X) = E[X^2] - (E[X])^2, \quad SD(X) = \sqrt{Var(X)}$$

Definition:

$$Var(X) = E[(X - \mu)^2]$$

Properties of $E[X]$ for continuous random variable

- Recall that expected value of the discrete random variable X with pmf $p(x)$,

$$E[g(X)] = \sum_{\text{all } x} g(x)p(x)$$

Properties of $E[X]$ for continuous random variable

- Recall that expected value of the discrete random variable X with pmf $p(x)$,

$$E[g(X)] = \sum_{\text{all } x} g(x)p(x)$$

- Similarly, expected value of a continuous random variable X with pdf $f(x)$ as

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Properties of $E[X]$ for continuous random variable

Similarly as discrete random variable, for continuous random variable X and Y

- Sum of two random variable's

$$E[X+Y] = E[X] + E[Y]$$

- If a and b are constants, then

$$E[aX+b] = aE[X] + b$$

- If a and b are constants, then

$$Var(aX+b) = a^2 Var(X)$$

Find the mean and variance of the continuous random variable X, whose pdf is

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean and variance of the continuous random variable X, whose pdf is

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{0}^{1} x \cdot 2x dx = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3} \end{aligned}$$

Find the mean and variance of the continuous random variable X, whose pdf is

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^1 x \cdot 2x dx = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3} \end{aligned}$$

$$Var[X] = E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - (2/3)^2$$

$$= \int_0^1 x^2 \cdot 2x dx - 4/9 = \frac{2}{4} x^4 \Big|_0^1 - 4/9 = \frac{1}{18}$$

Outline

Uniform Distribution

Let's define a continuous probability distribution that has some constant value c between α and β where $\alpha < \beta$. What is the pdf?

$$f(x) = \begin{cases} c & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

Uniform Distribution

Let's define a continuous probability distribution that has some constant value c between α and β where $\alpha < \beta$. What is the pdf?

$$f(x) = \begin{cases} c & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{\alpha}^{\beta} cdx = c(\beta - \alpha) \Longrightarrow c = \frac{1}{\beta - \alpha}$$

Uniform Distribution

Let's define a continuous probability distribution that has some constant value c between α and β where $\alpha < \beta$. What is the pdf?

$$f(x) = \begin{cases} c & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{\alpha}^{\beta} cdx = c(\beta - \alpha) \Longrightarrow c = \frac{1}{\beta - \alpha}$$

Definition

A continuous random variable X has a Uniform distribution on the interval (α, β) if its pdf is

$$X \sim \mathit{Unif}(\alpha, \beta) \Longleftrightarrow f(x) = \frac{1}{\beta - \alpha} \cdot \mathbf{1}_{(\alpha,\beta)}(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

Cdf of Uniform distribution

For any $x \in (\alpha, \beta)$,

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{\alpha}^{x} \frac{1}{\beta - \alpha}\ dt = \frac{x - \alpha}{\beta - \alpha}$$

Cdf of Uniform distribution

For any $x \in (\alpha, \beta)$,

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{\alpha}^{x} \frac{1}{\beta - \alpha}\ dt = \frac{x - \alpha}{\beta - \alpha}$$

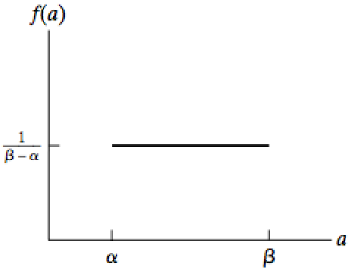
- $X \sim \text{Unif}(\alpha, \beta)$, then its cdf is $F(x) = \begin{cases} 0 & \text{if } x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 1 & \text{if } x \geq \beta \end{cases}$

Cdf of Uniform distribution

For any $x \in (\alpha, \beta)$,

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{\alpha}^{x} \frac{1}{\beta - \alpha}\ dt = \frac{x - \alpha}{\beta - \alpha}$$

- $X \sim$ Unif(α, β), then its cdf is $F(x) = \begin{cases} 0 & \text{if } x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 1 & \text{if } x \geq \beta \end{cases}$



Mean of a Uniform distribution

- (required) Expected value of $X \sim \text{Unif}(\alpha, \beta)$ is

$$E[X] = \frac{\alpha + \beta}{2}$$

Mean of a Uniform distribution

- (required) Expected value of $X \sim \text{Unif}(\alpha, \beta)$ is

$$E[X] = \frac{\alpha + \beta}{2}$$

$$\begin{aligned} E[X] &= \int_a^\beta \frac{x}{\beta - \alpha} dx = \left. \frac{x^2}{2(\beta - \alpha)} \right|_\alpha^\beta \\ &= \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} = \frac{\alpha + \beta}{2} \end{aligned}$$

Variance of a Uniform distribution

- (required) Variance of $X \sim \text{Unif}(\alpha, \beta)$ is

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$

Variance of a Uniform distribution

- (required) Variance of $X \sim \text{Unif}(\alpha, \beta)$ is

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$

$$E[X^2] = \int_a^\beta \frac{x^2}{\beta - \alpha} dx = \left. \frac{x^3}{3(\beta - \alpha)} \right|_\alpha^\beta$$

$$= \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \beta\alpha + \alpha^2}{3}$$

Variance of a Uniform distribution

- (required) Variance of $X \sim \text{Unif}(\alpha, \beta)$ is

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$

$$\begin{aligned} E[X^2] &= \int_a^\beta \frac{x^2}{\beta - \alpha} dx = \frac{x^3}{3(\beta - \alpha)}\Big|_\alpha^\beta \\ &= \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \beta\alpha + \alpha^2}{3} \end{aligned}$$

$$\begin{aligned} Var(X) = E[X^2] - E[X]^2 &= \frac{\beta^2 + \beta\alpha + \alpha^2}{3} - \frac{(\beta + \alpha)^2}{4} \\ &= \frac{4b^2 + 4\beta\alpha + 4\alpha^2}{12} - \frac{3b^2 + 6\alpha\beta + 3\alpha^2}{12} \\ &= \frac{\beta^2 - 2\beta\alpha + \alpha^2}{12} = \frac{(\beta - \alpha)^2}{12} \end{aligned}$$

Uniform distribution: probability calculation

If $X \sim \text{Unif}(\alpha, \beta)$, then $P(X \in B) = \dfrac{\text{length}(B)}{\beta - \alpha}$

Uniform distribution: probability calculation

If $X \sim \mathrm{Unif}(\alpha, \beta)$, then $P(X \in B) = \dfrac{\mathrm{length}(B)}{\beta - \alpha}$

Example: I want to take bus to the airport this afternoon. The bus arrives at my nearest bus stop at 10-minute intervals starting at 4pm. Suppose I arrive the bus stop at a time uniformly distributed between 4pm to 5pm. What's the probability I will wait more than 7 minutes?

Uniform distribution: probability calculation

If $X \sim \text{Unif}(\alpha, \beta)$, then $P(X \in B) = \dfrac{\text{length}(B)}{\beta - \alpha}$

Example: I want to take bus to the airport this afternoon. The bus arrives at my nearest bus stop at 10-minute intervals starting at 4pm. Suppose I arrive the bus stop at a time uniformly distributed between 4pm to 5pm. What's the probability I will wait more than 7 minutes?

Let X denote the time I arrive the bus stop.

$$X \sim \text{Unif}(0, 60)$$

Uniform distribution: probability calculation

If $X \sim \text{Unif}(\alpha, \beta)$, then $P(X \in B) = \dfrac{\text{length}(B)}{\beta - \alpha}$

Example: I want to take bus to the airport this afternoon. The bus arrives at my nearest bus stop at 10-minute intervals starting at 4pm. Suppose I arrive the bus stop at a time uniformly distributed between 4pm to 5pm. What's the probability I will wait more than 7 minutes?

Let X denote the time I arrive the bus stop.

$$X \sim \text{Unif}(0, 60)$$

The bus arrives at time $t = 0, 10, 20, 30, 40, 50, 60$.
So the intervals of my arriving time such that I will wait more than 7 min:

$$B = (0, 3) \cup (10, 13) \cup (20, 23) \cup (30, 33) \cup (40, 43) \cup (50, 53)$$

Uniform distribution: probability calculation

If $X \sim \text{Unif}(\alpha, \beta)$, then $P(X \in B) = \dfrac{\text{length}(B)}{\beta - \alpha}$

Example: I want to take bus to the airport this afternoon. The bus arrives at my nearest bus stop at 10-minute intervals starting at 4pm. Suppose I arrive the bus stop at a time uniformly distributed between 4pm to 5pm. What's the probability I will wait more than 7 minutes?

Let X denote the time I arrive the bus stop.

$$X \sim \text{Unif}(0, 60)$$

The bus arrives at time $t = 0, 10, 20, 30, 40, 50, 60$.
So the intervals of my arriving time such that I will wait more than 7 min:

$$B = (0,3) \cup (10,13) \cup (20,23) \cup (30,33) \cup (40,43) \cup (50,53)$$

$$P(X \in B) = (3 \times 6)/60 = 0.3$$

Recap

Expectation for continuous random variable X and a function of it $g(X)$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Recap

Expectation for continuous random variable X and a function of it $g(X)$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Uniform distribution $X \sim \text{Unif}(\alpha, \beta)$

- Pdf
$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

- Cdf
$$F(x) = \begin{cases} 0 & \text{if } x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 1 & \text{if } x \geq \beta \end{cases}$$

- Mean and variance
$$E[X] = \frac{\alpha + \beta}{2}, \quad Var(X) = \frac{(\beta - \alpha)^2}{12}$$