Week 5: Continuous Random Variables and Probability Distributions

MATH697

Sahir Bhatnagar

October 3, 2017

McGill University

Introduction

- In the previous section, we considered discrete random variables X for which

$$P(X = x) > 0$$

for certain values of x.

- In the previous section, we considered discrete random variables X for which
$$P(X = x) > 0$$
for certain values of x.
- However, for some random variables X, such as one having the uniform distribution, we have
$$P(X = x) = 0, \quad \forall \quad x \in \mathbb{R}$$

- In the previous section, we considered discrete random variables X for which
$$P(X = x) > 0$$
for certain values of x.
- However, for some random variables X, such as one having the uniform distribution, we have
$$P(X = x) = 0, \quad \forall \quad x \in \mathbb{R}$$

Definition 1 (Continuous Random Variable)

A random variable is continuous if

$$P(X = x) = 0, \quad \forall \quad x \in \mathbb{R}$$

Example

Example 2 (The Uniform Distribution)

Consider a random variable X whose distribution is the uniform distribution on $[0, 1]$:

$$P(a \leq X \leq b) = b - a, \quad 0 \leq a \leq b \leq 1$$

with $P(X < 0) = P(X > 1) = 0$. We write this as $X \sim \mathrm{Uniform}(0, 1)$.

Example

Example 2 (The Uniform Distribution)

Consider a random variable X whose distribution is the uniform distribution on $[0, 1]$:

$$P(a \leq X \leq b) = b - a, \quad 0 \leq a \leq b \leq 1$$

with $P(X < 0) = P(X > 1) = 0$. We write this as $X \sim \mathrm{Uniform}(0, 1)$.

$$P\left(\frac{1}{2} \leq X \leq \frac{3}{4}\right) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$$

Example 2 (The Uniform Distribution)

Consider a random variable X whose distribution is the uniform distribution on $[0, 1]$:

$$P(a \leq X \leq b) = b - a, \quad 0 \leq a \leq b \leq 1$$

with $P(X < 0) = P(X > 1) = 0$. We write this as $X \sim \mathrm{Uniform}(0, 1)$.

$$P\left(\frac{1}{2} \leq X \leq \frac{3}{4}\right) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$$

$$P\left(X \geq \frac{2}{3}\right) = P\left(\frac{2}{3} \leq X \leq 1\right) + P(X > 1) = \left(1 - \frac{2}{3}\right) + 0$$

Example cont

Example 3 (The Uniform Distribution)

$X \sim \mathrm{Uniform}(0, 1)$. Setting $a = b = x$ we see in particular that

$$P\left(x \leq X \leq x\right) = x - x = 0, \quad \forall \quad x \in \mathbb{R}$$

Example cont

Example 3 (The Uniform Distribution)

$X \sim \mathrm{Uniform}(0, 1)$. Setting $a = b = x$ we see in particular that

$$P\left(x \leq X \leq x\right) = x - x = 0, \quad \forall \quad x \in \mathbb{R}$$

Thus, the uniform distribution is an example of a continuous distribution.

Continous CDF and Probability Density Function (PDF)

Continous CDF

The probability distribution of a *continuous* random variable X is defined by the continuous **cumulative distribution function** or **c.d.f.**, F_X, specified by

$$F_X(x) = P[X \leq x] \qquad \text{for all } x \in \mathbb{X}$$

note: this is an identical definition to the discrete case.

Continous CDF

The probability distribution of a *continuous* random variable X is defined by the continuous **cumulative distribution function** or **c.d.f.**, F_X, specified by

$$F_X(x) = P[X \leq x] \qquad \text{for all } x \in \mathbb{X}$$

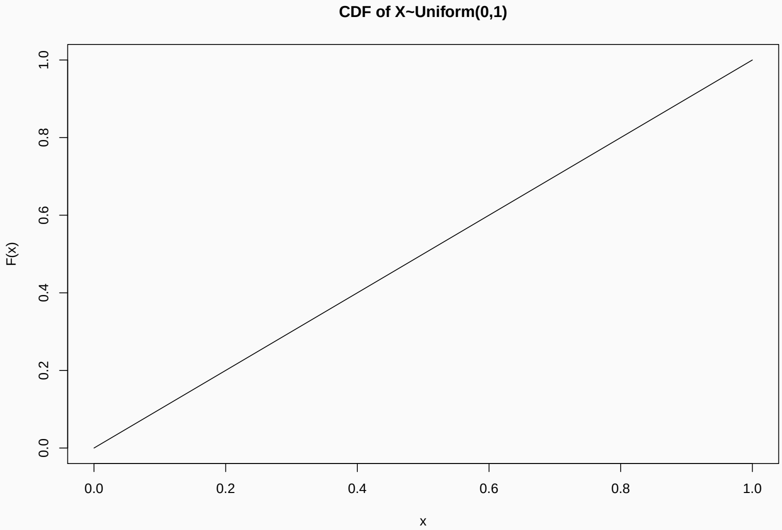
note: this is an identical definition to the discrete case.

The continuous cdf F_X must exhibit the same properties as for the discrete cdf, except that

(iii) $\lim_{h \to 0} F_X(x+h) = F_X(x)$ [i.e. F_X is continuous]

Uniform(0,1) Distribution

```
curve(punif(x), ylab = "F(x)",  main = "CDF of X~Uniform(0,1)")
```



Continous PDF

Definition 4 (Continuous Random Variable)

Let X be a random variable with distribution function $F(x)$. If there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that

$$F_X(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}$$

then X is called a continous random variable with density function f.

Continous PDF

Definition 4 (Continuous Random Variable)

Let X be a random variable with distribution function $F(x)$. If there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that

$$F_X(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}$$

then X is called a continous random variable with density function f. By the fundamental theorem of calculus, we also have

$$F'(x) = f(x), \quad \forall \quad x$$

Continous PDF

Proposition 5 (Continuous PDF)

f has the following properties

1. $f(x) \geq 0$ *for all* $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$

Proof: *on board*

Proposition 6 (CDF from PDF)

Let X be a continuous random variable with density function f

1. *If $a < b$, then*

$$P(a < X \le b) = \int_a^b f(x)dx$$

   Notat that this is the area under the curve of f between a and b.

   More generally, we have

$$P(X \in A) = \int_A f(x)dx, \quad for\ any\ A \subset \mathbb{R}$$

Continous PDF

Proposition 6 (CDF from PDF)

Let X be a continuous random variable with density function f

1. *If $a < b$, then*

$$P(a < X \leq b) = \int_a^b f(x)dx$$

Notat that this is the area under the curve of f between a and b.

More generally, we have

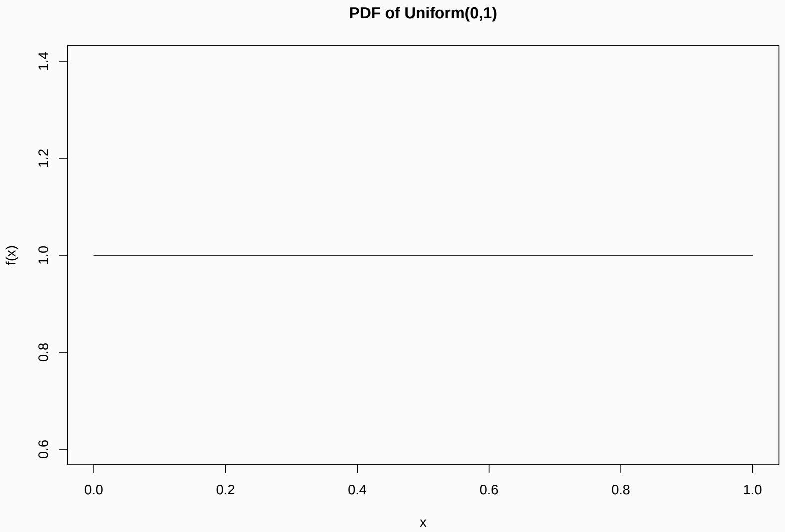
$$P(X \in A) = \int_A f(x)dx, \quad for\ any\ A \subset \mathbb{R}$$

2. *$P(X = x) = 0$, for every $x \in \mathbb{R}$*

Proof: *on board*

Example

```r
# X ~ Uniform(0,1)
curve(dunif(x), xlab = "x", ylab = "f(x)", main = "PDF of Uniform(0,1)")
```



Continuous PDF

Remark: Because of part (2) on the previous slide, we can say that

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

Relationship between CDF and PDF

The **probability density function**, or **pdf**, f_X, is defined by

$$f_X(x) = \frac{d}{dx}\left\{F_X(x)\right\}$$

so that, by a fundamental calculus result,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

Relationship between CDF and PDF

The **probability density function**, or **pdf**, f_X, is defined by

$$f_X(x) = \frac{d}{dx}\left\{F_X(x)\right\}$$

so that, by a fundamental calculus result,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

In the continuous case, we calculate F_X from f_X by **integration**, and f_X from F_X by **differentiation**

Remarks about PDF and CDF

- We must use F_X to specify the probability distribution initially, although it is often easier to think of the shape of the distribution via the pdf f_X. Any function that satisfies the properties for a pdf can be used to construct a probability distribution.

Remarks about PDF and CDF

- We must use F_X to specify the probability distribution initially, although it is often easier to think of the shape of the distribution via the pdf f_X. Any function that satisfies the properties for a pdf can be used to construct a probability distribution.
- Note that, for a continuous random variable

$$f_X(x) \neq P[X = x].$$

Special Continous Probability Distributions

1. Uniform Distribution

Definition 7 (The continous uniform distribution)

A model with **constant** probability density on a region,

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

the cumulative distribution function (cdf) is also straightforward

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

2. The Exponential Distribution

Definition 8 (The Exponential(λ) Distribution)

A **continuous** waiting-time model

$$f_X(x) = \lambda e^{-\lambda x} \qquad x \in \mathbb{R}^+$$

The cdf for the exponential distribution can be calculated easily;

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x} \qquad x \geq 0.$$

and note that

$$P[\, X > x\,] = 1 - P[\, X \leq x\,] = 1 - F_X(x) = e^{-\lambda x}$$

- This is the parametrization used in R (see `?dexp`)

Memoryless Property of the Exponential

The exponential distribution can be used to model lifetimes as it shares the memoryless property of the geometric. If $X \sim Exponential(\lambda)$, then for $s > t \geq 0$

$$P(X > s | X > t) = P(X > s - t)$$

Memoryless Property of the Exponential

The exponential distribution can be used to model lifetimes as it shares the memoryless property of the geometric. If $X \sim Exponential(\lambda)$, then for $s > t \geq 0$

$$P(X > s | X > t) = P(X > s - t)$$

- Tossing a fair coin is an example that is memoryless. Every time you toss the coin, you have a 50% chance of it coming up heads. It doesn't matter whether or not the last 5 times you tossed the coin it came up consistently tails; the probability of heads in the next throw is always going to be 0.5

Memoryless Property of the Exponential

The exponential distribution can be used to model lifetimes as it shares the memoryless property of the geometric. If $X \sim Exponential(\lambda)$, then for $s > t \geq 0$

$$P(X > s | X > t) = P(X > s - t)$$

- Tossing a fair coin is an example that is memoryless. Every time you toss the coin, you have a 50% chance of it coming up heads. It doesn't matter whether or not the last 5 times you tossed the coin it came up consistently tails; the probability of heads in the next throw is always going to be 0.5
- Time until car failure is not memoryless.

$$P(7 \text{ years} < fail < 10 \text{ years}) \neq P(3 \text{ years} < fail < 6 \text{ years})$$

Memoryless Property of the Exponential (proof)

$$
\begin{aligned}
P(X > s | X > t) &= \frac{P(X > s \cap X > t)}{P(X > t)} \\
&= \frac{P(X > s)}{P(X > t)} \\
&= \frac{\int_s^\infty \lambda e^{-\lambda x} dx}{\int_t^\infty \lambda e^{-\lambda x} dx} \\
&= \frac{e^{-\lambda s}}{e^{-\lambda t}} \\
&= e^{-\lambda(s-t)} \\
&= 1 - F_X(s-t) \\
&= P(X > s - t)
\end{aligned}
$$

3. The Gamma Distribution

Gamma Function

The *gamma function* is defined by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \qquad \alpha > 0 \tag{1}$$

It turns out that

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$$

and that

- $\Gamma(\alpha) = (\alpha - 1)! \rightarrow$ if α is a positive integer

The *gamma function* is defined by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt, \qquad \alpha > 0 \tag{1}$$

It turns out that

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$$

and that

- $\Gamma(\alpha) = (\alpha - 1)! \rightarrow$ if α is a positive integer
- $\Gamma(1/2) = \sqrt{\pi}$

Gamma Distribution

We can use the gamma function to define the density of the Gamma(α, β) distribution.

Definition 9 (The Gamma(α, β) Distribution)

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \qquad 0 < x < \infty, \quad \alpha > 0, \quad \beta > 0, \tag{2}$$

where $\Gamma(\alpha)$ is defined in Equation (1) on the previous slide.

- A random variable X having density function f given by (2) is said to have the Gamma(α, β) distribution.

We can use the gamma function to define the density of the Gamma(α, β) distribution.

Definition 9 (The Gamma(α, β) Distribution)

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \qquad 0 < x < \infty, \quad \alpha > 0, \quad \beta > 0, \tag{2}$$

where $\Gamma(\alpha)$ is defined in Equation (1) on the previous slide.

- A random variable X having density function f given by (2) is said to have the Gamma(α, β) distribution.
- We write this as $X \sim \mathrm{Gamma}(\alpha, \beta)$

We can use the gamma function to define the density of the $\text{Gamma}(\alpha, \beta)$ distribution.

Definition 9 (The $\text{Gamma}(\alpha, \beta)$ Distribution)

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \qquad 0 < x < \infty, \quad \alpha > 0, \quad \beta > 0, \tag{2}$$

where $\Gamma(\alpha)$ is defined in Equation (1) on the previous slide.

- A random variable X having density function f given by (2) is said to have the $\text{Gamma}(\alpha, \beta)$ distribution.
- We write this as $X \sim \text{Gamma}(\alpha, \beta)$
- This is the parametrization used in `R` (see `?dgamma`)

Example 10 (The Gamma Distribution)

Verify that Equation (2) is really a density function.

Proof: on board

Gamma Distribution Parameters

- The parameter α is known as the shape parameter $\longrightarrow$ influences the peakedness of the distribution

Gamma Distribution Parameters

- The parameter α is known as the shape parameter $\longrightarrow$ influences the peakedness of the distribution
- The parameter β is called the scale parameter $\longrightarrow$ influences the spread of the distribution

Gamma Distribution Parameters

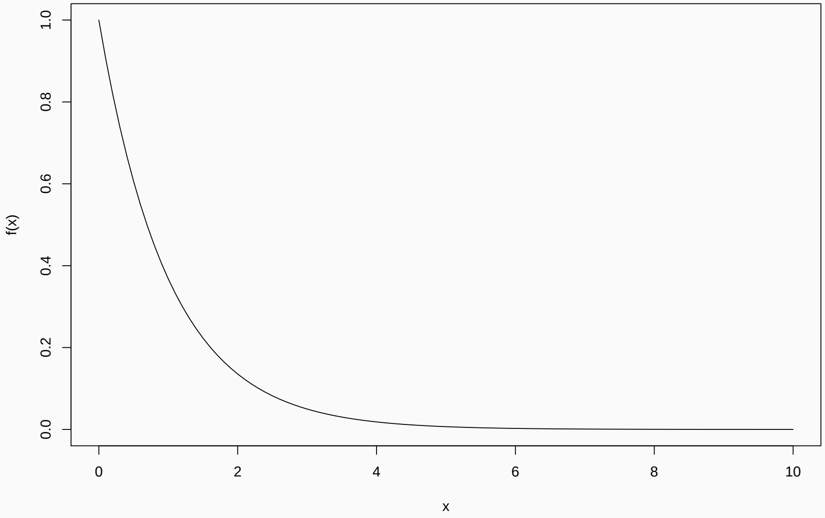
- The parameter α is known as the shape parameter $\longrightarrow$ influences the peakedness of the distribution
- The parameter β is called the scale parameter $\longrightarrow$ influences the spread of the distribution
- $\mathrm{Gamma}(1, \beta) \equiv \mathrm{Exponential}(\beta)$

Gamma Distribution Parameters

- The parameter α is known as the shape parameter $\longrightarrow$ influences the peakedness of the distribution
- The parameter β is called the scale parameter $\longrightarrow$ influences the spread of the distribution
- $\text{Gamma}(1, \beta) \equiv \text{Exponential}(\beta)$
- $\text{Gamma}(p/2, 2) \equiv \chi^2_{(p)}, p = \{0, 1, 2, 3, \ldots\}$

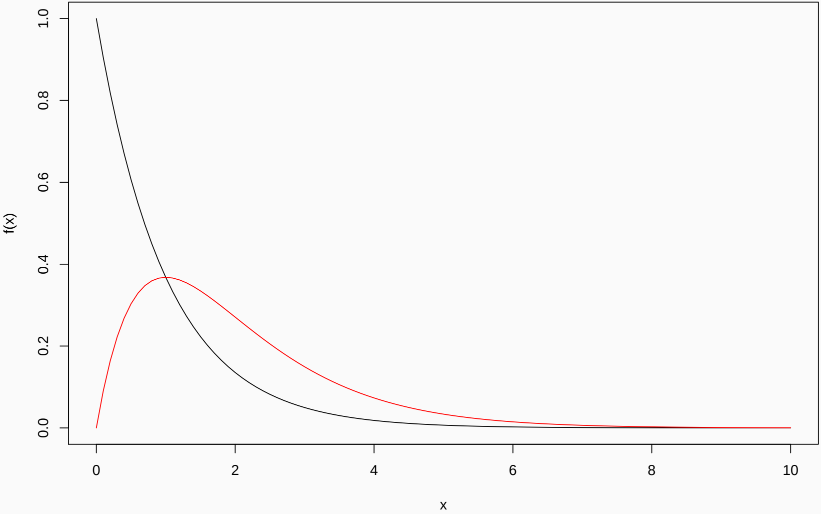
Gamma Distribution Parameters

```
curve(dgamma(x, shape = 1, scale = 1), from = 0, to = 10, ylab = "f(x)")
```



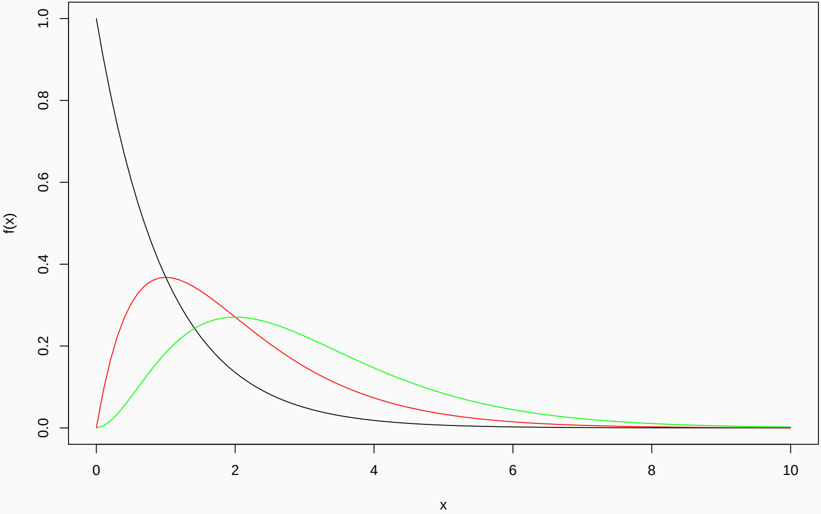
Gamma Distribution Parameters

```
curve(dgamma(x, shape = 2, scale = 1), add = TRUE, col = "red")
```



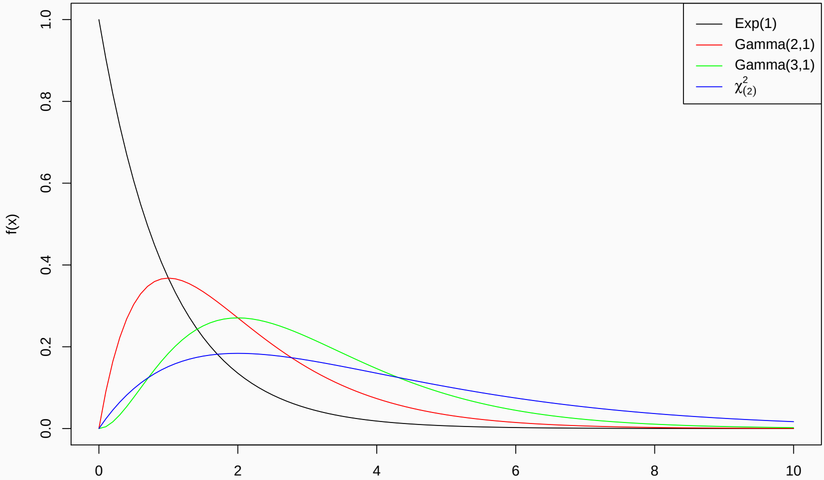
Gamma Distribution Parameters

```r
curve(dgamma(x, shape = 3, scale = 1), add = TRUE, col = "green")
```



Gamma Distribution Parameters

```
curve(dgamma(x, shape = 4/2, scale = 2), add = TRUE, col = "blue")
legend("topright",
       legend = c("Exp(1)", "Gamma(2,1)", "Gamma(3,1)",expression(chi[(2)]^2)),
       col = c("black","red","green","blue"), lty = 1)
```



Gamma Distribution (alternative definition)

Some books and software packages replace β with $1/\beta$:

Definition 11 (The Gamma(α, β) Distribution)

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \qquad x \in \mathbb{R}^+$$

Relation between Exponential and Gamma Distribution

The Gamma distribution is another **continuous** waiting-time model. It can be shown the sum of i.i.d. Exponential random variables has a Gamma distribution, that is, if $X_1, X_2, \ldots, X_n$ are independent and identically distributed $Exponential(\lambda)$ random variables, then

$$X = \sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$$