Probability 1

Chapter 05 : Continuous Random Variables - Part 1

Dr. Daniel Flores-Agreda,

(based on the notes of Prof. Davide La Vecchia)

Spring Semester 2021

Objectives

- Present the Exponential distribution
- Understand the consequences of the transformation of Random Variables.

Outline

1. Exponential distribution
2. Variable Transformation
   - Transformation of Discrete Random Variables
   - Transformation through the CDF
   - Transformation of Continuous Random Variables through the PDF

Outline

Exponential distribution

Definition

Let X be a continuous random variable, having the following characteristics:

- X is defined on the **positive real numbers** $(0; \infty)$ i.e. $\mathbb{R}^+$;
- the PDF and CDF are

$$f_X(x) = \lambda \exp\{-\lambda x\}, \lambda > 0; \quad F_X(x) = 1 - \exp(-\lambda x);$$

then we say that X has an **exponential distribution**.

We write $X \sim \text{Exp}(\lambda)$.

Exponential distribution

Remark (Expectation and Variance)

For $X \sim Exp(\lambda)$ *we have that:*

$$E[X] = \frac{1}{\lambda} \quad \text{and} \quad Var(X) = \frac{1}{\lambda^2}. \tag{1}$$

Exponential distribution

Remark (Expectation and Variance)

For $X \sim Exp(\lambda)$ we have that:

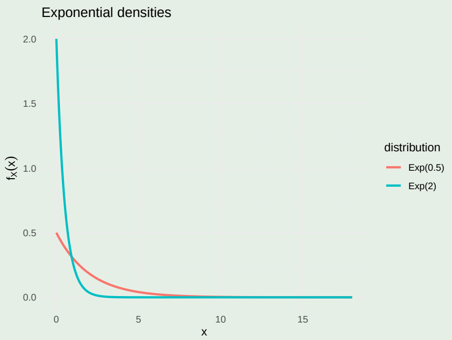
$$E[X] = \frac{1}{\lambda} \quad \text{and} \quad Var(X) = \frac{1}{\lambda^2}. \tag{1}$$

Remark (Applications and Properties)

X is typically applied to model the time until an event occurs, *when events are always occurring at a rate $\lambda > 0$.*

The sum of independent exponential *random variables has a* Gamma distribution *(see the exercises).*

Example (A graphical illustration)



Example (Illustration of use (Spoiler of the Exercises!))

The lifetime X in years of a television follows an exponential law with density:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

1. Compute the cumulative distribution function $F(x)$.
2. Compute the α-quantile $q_\alpha = F^{-1}(\alpha)$.
3. The expected life of your television is 8 years. What is the probability that the lifetime of your television is more than 8 years ? Evaluate the median.
4. Compute the variance of X for any λ.

Exponential distribution

Unlike the Normal, the CDF of an exponential has a closed-form expression

Example (CDF of Exponential)

Let $X \sim \text{Exp}(\lambda)$, with $\lambda = 0.5$. Thus

$$f_X(x) = \begin{cases} 0.5 \exp(-0.5x) & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, find the CDF.

Exponential distribution

Unlike the Normal, the CDF of an exponential has a closed-form expression

Example (CDF of Exponential)

Let $X \sim \text{Exp}(\lambda)$, with $\lambda = 0.5$. Thus

$$f_X(x) = \begin{cases} 0.5\exp(-0.5x) & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, find the CDF.

For $x > 0$, we have

$$\begin{aligned} F_X(x) &= \int_0^x f_X(u)du \\ &= 0.5\Big(-2\exp(-0.5u)\Big)\Big|_{u=0}^{u=x} \\ &= 0.5(-2\exp(-0.5x) + 2\exp(0)) \\ &= 1 - \exp(-0.5x) \end{aligned}$$

Exponential distribution

Unlike the Normal, the CDF of an exponential has a closed-form expression

Example (CDF of Exponential)

Let $X \sim \text{Exp}(\lambda)$, with $\lambda = 0.5$. Thus

$$f_X(x) = \begin{cases} 0.5\exp(-0.5x) & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, find the CDF.

For $x > 0$, we have

$$\begin{aligned} F_X(x) &= \int_0^x f_X(u)du \\ &= 0.5\Big(-2\exp(-0.5u)\Big)\Big|_{u=0}^{u=x} \\ &= 0.5(-2\exp(-0.5x) + 2\exp(0)) \\ &= 1 - \exp(-0.5x) \end{aligned}$$

so, finally,

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 1 - \exp(-0.5x) & x > 0 \end{cases}$$

Outline

Variable Transformation

Remark (Problem)

Consider a random variable X *and suppose* we are interested in $Y = \psi(X)$, *where* ψ *is a* one to one function

Variable Transformation

Remark (Problem)

Consider a random variable X *and suppose* we are interested in $Y = \psi(X)$, *where* ψ *is a* one to one function

Definition

A **function** $\psi(x)$ **is one to one** (1-to-1) if there are no two numbers, x_1, x_2 in the domain of ψ such that $\psi(x_1) = \psi(x_2)$ but $x_1 \neq x_2$.

Variable Transformation

Remark (Problem)

Consider a random variable X and suppose we are interested in $Y = \psi(X)$, *where ψ is a* one to one function

Definition

A **function** $\psi(x)$ **is one to one** (1-to-1) if there are no two numbers, x_1, x_2 in the domain of ψ such that $\psi(x_1) = \psi(x_2)$ but $x_1 \neq x_2$.

Remark

A sufficient condition for $\psi(x)$ to be 1-to-1 is that it be monotonically increasing (or decreasing) in x.

Variable Transformation

Remark (Problem)

Consider a random variable X and suppose we are interested in $Y = \psi(X)$, *where ψ is a* one to one function

Definition

A **function** $\psi(x)$ **is one to one** (1-to-1) if there are no two numbers, x_1, x_2 in the domain of ψ such that $\psi(x_1) = \psi(x_2)$ but $x_1 \neq x_2$.

Remark

A sufficient condition for $\psi(x)$ to be 1-to-1 is that it be monotonically increasing (or decreasing) in x. Note that the inverse *of a 1-to-1 function $y = \psi(x)$ is a*

1-to-1 function $\psi^{-1}(y)$ such that

$$\psi^{-1}(\psi(x)) = x \textit{ and } \psi\left(\psi^{-1}(y)\right) = y.$$

To transform X to Y, we need to consider all the values x that X can take
We first transform x into values $y = \psi(x)$

Outline

- To transform a discrete random variable X, into the random variable $Y = \psi(X)$, we transfer the probabilities for **each** x to the values $y = \psi(x)$:

Probability function for X

X	$P(\{X = x_i\}) = p_i$
x_1	p_1
x_2	p_2
x_3	p_3
$\vdots$	$\vdots$
x_n	p_n

$\Rightarrow$

Probability function for X

Y	$P(\{X = x_i\}) = p_i$
$\psi(x_1)$	p_1
$\psi(x_2)$	p_2
$\psi(x_3)$	p_3
$\vdots$	$\vdots$
$\psi(x_n)$	p_n

- To transform a discrete random variable X, into the random variable $Y = \psi(X)$, we transfer the probabilities for **each** x to the values $y = \psi(x)$:

Probability function for X

X	$P(\{X = x_i\}) = p_i$
x_1	p_1
x_2	p_2
x_3	p_3
$\vdots$	$\vdots$
x_n	p_n

$\Rightarrow$

Probability function for X

Y	$P(\{X = x_i\}) = p_i$
$\psi(x_1)$	p_1
$\psi(x_2)$	p_2
$\psi(x_3)$	p_3
$\vdots$	$\vdots$
$\psi(x_n)$	p_n

- Note that this is equivalent to applying the function $\psi(\cdot)$ inside the probability statements:

$$\begin{aligned} P(\{X = x_i\}) &= P(\{\psi(X) = \psi(x_i)\}) \\ &= P(\{Y = y_i\}) \\ &= p_i \end{aligned}$$

Example (option pricing)

Let us imagine that we are tossing a balanced coin ($p = 1/2$), and when we get a "Head" (H) the stock price moves up of a factor u, but when we get a "Tail" (T) the price moves down of a factor d. We denote the price at time t_1 by $S_1(H) = uS_0$ if the toss results in head (H), and by $S_1(T) = dS_0$ if it results in tail (T). After the second toss, the price will be one of:

$$S_2(HH) = uS_1(H) = u^2S_0, \quad S_2(HT) = dS_1(H) = duS_0,$$

$$S_2(TH) = uS_1(T) = udS_0, \quad S_2(TT) = dS_1(T) = d^2S_0.$$

Indeed, after two tosses, there are four possible coin sequences,

$$\{HH, HT, TH, TT\}$$

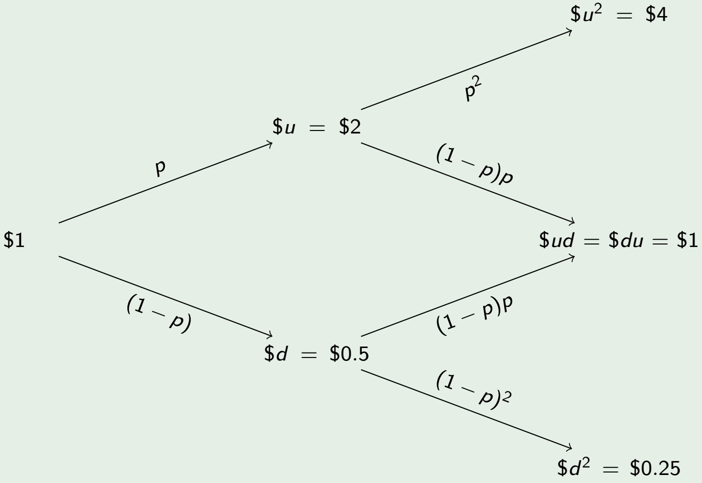
although not all of them result in different stock prices at time t_2.

Variable Transformation

Transformation of Discrete Random Variables

Example (continued)

Let us set $S_0 = 1$, $u = 2$ and $d = 1/2$: we represent the price evolution by a tree:



Example (continued)

Now consider an European option call with maturity t_2 and strike price $K = 0.5$, whose random pay-off at t_2 is $C = \max(0; S_2 - 0.5)$. Thus,

$$C(HH) = \max(0; 4 - 0.5) = \$3.5 \quad C(HT) = \max(0; 1 - 0.5) = \$0.5$$
$$C(TH) = \max(0; 1 - 0.5) = \$0.5 \quad C(TT) = \max(0; 0.25 - 0.5) = \$0.$$

Thus at maturity t_2 we have

Probability function for S_2

S_2	$P(\{X = x_i\}) = p_i$
$\$u^2$	p^2
$\$ud$	$2p(1-p)$
$\$d^2$	$(1-p)^2$

$\Rightarrow$

Probability function for C

C	$P(\{C = c_i\}) = p_i$
$\$3.5$	p^2
$\$0.5$	$2p(1-p)$
$\$0$	$(1-p)^2$

Since $ud = du$ the corresponding values of S_2 and C can be aggregated, without loss of info.

Outline

- We can use the same logic for CDF probabilities, whether the random variables are **discrete or continuous**

- We can use the same logic for CDF probabilities, whether the random variables are **discrete or continuous**
- Let $Y = \psi(X)$ with $\psi(x)$ **1-to-1 and monotone increasing**. Then

$$
\begin{aligned}
F_Y(y) &= P(\{Y \leq y\}) \\
&= P(\{\psi(X) \leq y\}) = P\left(\{X \leq \psi^{-1}(y)\}\right) \\
&= F_X\left(\psi^{-1}(y)\right)
\end{aligned}
$$

- We can use the same logic for CDF probabilities, whether the random variables are **discrete or continuous**
- Let $Y = \psi(X)$ with $\psi(x)$ **1-to-1 and monotone increasing**. Then

$$\begin{aligned} F_Y(y) &= P(\{Y \leq y\}) \\ &= P(\{\psi(X) \leq y\}) = P\left(\{X \leq \psi^{-1}(y)\}\right) \\ &= F_X\left(\psi^{-1}(y)\right) \end{aligned}$$

Example

Let $Y = \psi(X) = \exp X$ where $X \sim F_X$ on all values $x \in \mathbb{R}$

$$\begin{aligned} F_Y(y) &= P(\{Y \leq y\}) \\ &= P(\{\exp X \leq y\}) = P(\{X \leq \ln(y)\}) \\ &= F_X(\ln(y)) \text{ only for } y > 0. \end{aligned}$$

- **Monotone decreasing functions** work in a similar way, but require **changing the sense of the inequality**.
- Let $Y = \psi(X)$ with $\psi(x)$ 1-to-1 and **monotone decreasing**. Then

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

- **Monotone decreasing functions** work in a similar way, but require **changing the sense of the inequality**.
- Let $Y = \psi(X)$ with $\psi(x)$ 1-to-1 and **monotone decreasing**. Then

$$\begin{aligned} F_Y(y) &= P(\{Y \le y\}) \\ &= P(\{\psi(X) \le y\}) = P\left(\{X \ge \psi^{-1}(y)\}\right) \\ &= 1 - F_X\left(\psi^{-1}(y)\right) \end{aligned}$$

Example

Example: let $Y = \psi(X) = -\exp X$ where $X \sim F_X$ on all values $x \in \mathbb{R}$

$$\begin{aligned} F_Y(y) &= P(\{Y \le y\}) = P\left(\{-\exp^X \le y\}\right) \\ &= P(\{\exp X \ge -y\}) = P(\{X \ge \ln(-y)\}) \\ &= 1 - F_X(\ln(-y)) \text{ only for } y < 0. \end{aligned}$$

Outline

1. Exponential distribution
2. Variable Transformation
   - Transformation of Discrete Random Variables
   - Transformation through the CDF
   - Transformation of Continuous Random Variables through the PDF

- For continuous random variables, if $\psi(x)$ 1-to-1 and monotone **increasing**, we have

$$F_Y(y) = F_X\left(\psi^{-1}(y)\right)$$

- For continuous random variables, if $\psi(x)$ 1-to-1 and monotone **increasing**, we have

$$F_Y(y) = F_X\left(\psi^{-1}(y)\right)$$

- Notice this implies that the pdf of $Y = \psi(X)$ must satisfy

$$\begin{aligned}
f_Y(y) &= \frac{dF_Y(y)}{dy} = \frac{dF_X\left(\psi^{-1}(y)\right)}{dy} \\
&= \frac{dF_X(x)}{dx} \times \frac{d\psi^{-1}(y)}{dy} \qquad \text{(chain rule)} \\
&= f_X(x) \times \frac{d\psi^{-1}(y)}{dy} \qquad \text{(derivative of CDF (of } X\text{) is pdf)} \\
&= f_X\left(\psi^{-1}(y)\right) \times \frac{d\psi^{-1}(y)}{dy} \qquad \text{(substitute } x = \psi^{-1}(y)\text{)}
\end{aligned}$$

- What happens when $\psi(x)$ 1-to-1 and monotone **decreasing**? We have

$$F_Y(y) = 1 - F_X\left(\psi^{-1}(y)\right)$$

- What happens when $\psi\left(x\right)$ 1-to-1 and monotone **decreasing**? We have

$$F_Y\left(y\right) = 1 - F_X\left(\psi^{-1}\left(y\right)\right)$$

- So now the pdf of $Y = \phi\left(X\right)$ must satisfy

$$\begin{aligned} f_Y\left(y\right) &= \frac{dF_Y\left(y\right)}{dy} = -\frac{dF_X\left(\psi^{-1}\left(y\right)\right)}{dy} \\ &= -f_X\left(\psi^{-1}\left(y\right)\right) \times \frac{d\psi^{-1}\left(y\right)}{dy} \qquad \text{(same reasons as before)} \end{aligned}$$

- What happens when $\psi(x)$ 1-to-1 and monotone **decreasing**? We have

$$F_Y(y) = 1 - F_X\left(\psi^{-1}(y)\right)$$

- So now the pdf of $Y = \phi(X)$ must satisfy

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} = -\frac{dF_X\left(\psi^{-1}(y)\right)}{dy} \\ &= -f_X\left(\psi^{-1}(y)\right) \times \frac{d\psi^{-1}(y)}{dy} \qquad \text{(same reasons as before)} \end{aligned}$$

- but $\frac{d\psi^{-1}(y)}{dy} < 0$ since here $\psi(\cdot)$ is monotone decreasing, hence we can write

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

- What happens when $\psi(x)$ 1-to-1 and monotone **decreasing**? We have

$$F_Y(y) = 1 - F_X\left(\psi^{-1}(y)\right)$$

- So now the pdf of $Y = \phi(X)$ must satisfy

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} = -\frac{dF_X\left(\psi^{-1}(y)\right)}{dy} \\ &= -f_X\left(\psi^{-1}(y)\right) \times \frac{d\psi^{-1}(y)}{dy} \qquad \text{(same reasons as before)} \end{aligned}$$

- but $\frac{d\psi^{-1}(y)}{dy} < 0$ since here $\psi(\cdot)$ is monotone decreasing, hence we can write

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

- This expression (called Jacobian-formula) is valid for $\psi(x)$ 1-to-1 and monotone (whether increasing or decreasing)

Example

- So what is the pdf for the lognormal distribution?

Example

- So what is the pdf for the lognormal distribution?
- Recall that Y has a **lognormal distribution** when $\ln(Y) = X$ has a Normal distribution

Example

- So what is the pdf for the lognormal distribution?
- Recall that Y has a **lognormal distribution** when $\ln(Y) = X$ has a Normal distribution
- $\Rightarrow$ if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = \exp X \sim \textit{lognormal}(\mu, \sigma^2)$

Example

- So what is the pdf for the lognormal distribution?
- Recall that Y has a **lognormal distribution** when $\ln(Y) = X$ has a Normal distribution
- $\Rightarrow$ if $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $Y = \exp X \sim \textit{lognormal}(\mu, \sigma^2)$
  - Corresponding to $\psi(x) = \exp x$ and $\psi^{-1}(y) = \ln(y)$

Example

- So what is the pdf for the lognormal distribution?
- Recall that Y has a **lognormal distribution** when $\ln(Y) = X$ has a Normal distribution
- $\Rightarrow$ if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = \exp X \sim$ *lognormal*(μ, σ^2)
  - Corresponding to $\psi(x) = \exp x$ and $\psi^{-1}(y) = \ln(y)$
- The *pdf* of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

for any $-\infty < x < \infty$

Example

- So what is the pdf for the lognormal distribution?
- Recall that Y has a **lognormal distribution** when $\ln(Y) = X$ has a Normal distribution
- $\Rightarrow$ if $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $Y = \exp X \sim \textit{lognormal}(\mu, \sigma^2)$
  - Corresponding to $\psi(x) = \exp x$ and $\psi^{-1}(y) = \ln(y)$
- The *pdf* of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

  for any $-\infty < x < \infty$
- Using $\psi(x) = \exp x$ we know we'll have possible values for Y only on $0 < y < \infty$

Example (continued)

- We know that

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

Example (continued)

- We know that

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

- And since $\psi^{-1}(y) = \ln(y)$ then

$$\left|\frac{d\psi^{-1}(y)}{dy}\right| = \left|\frac{1}{y}\right|$$

Example (continued)

- We know that

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

- And since $\psi^{-1}(y) = \ln(y)$ then

$$\left|\frac{d\psi^{-1}(y)}{dy}\right| = \left|\frac{1}{y}\right|$$

- $\Rightarrow$ the *pdf* of Y is

$$f_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\ln(y) - \mu\right)^2\right\}$$

  for any $0 < y < \infty$

Example (continued)

- Both the Normal and the lognormal are characterized by only two parameters (μ and σ). The *median* of the lognormal distribution is $\exp \mu$, since

$$P(\{X \leq \mu\}) = 0.5,$$

and hence

$$\begin{aligned} 0.5 &= P(\{X \leq \mu\}) \\ &= P(\{\exp X \leq \exp \mu\}) \\ &= P(\{Y \leq \exp \mu\}). \end{aligned}$$

When X and Y are two random variables, **we should pay attention to their transformations**.

For instance, let us consider

$$X \sim \mathcal{N}(\mu, \sigma^2) \quad \text{and} \quad Y \sim Exp(\lambda).$$

Then, let's transform X and Y

- in a linear way: $Z = X + Y$. We know that

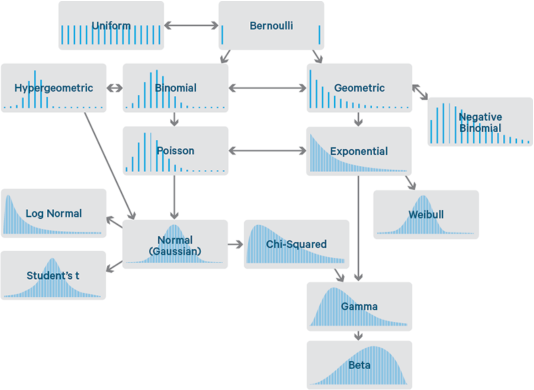
$$E[Z] = E[X + Y] = E[X] + E[Y]$$

- in a nonlinear way $W = X/Y$. One can show that

$$E[W] = E\left[\frac{X}{Y}\right] \neq \frac{E[X]}{E[Y]}.$$

Despite exotic names, the common distributions relate to each other in intuitive and interesting ways. Several follow naturally from the Bernoulli distribution, for example.

- *'Common probability distributions: the data scientist's crib sheet'* (goo.gl/NJRlXn):



Wrap-up

- The Exponential distribution helps to model *duration* data.

- To compute the probabilities of transformed Discrete Random Variables we proceed on the Probability Function:

$$P(X = x) = P(\psi(X) = \psi(x)) = P(Y = \psi(x))$$

- This principle applies to the inequalities in the CDF (give or take the sense of the monotonicity)

$$F_X(x) = P(X \leq x) = P(\psi(X) \leq \psi(x)) = P(Y \leq \psi(x)) = F_Y(\psi(x))$$
$$F_X(x) = P(X \leq x) = P(\psi(X) \geq \psi(x)) = P(Y \geq \psi(x)) = 1 - F_Y(\psi(x))$$

- The density of a transformed variable can be found with the formula:

$$f_Y(y) = f_X\left(\psi^{-1}(y)\right) \times \left|\frac{d\psi^{-1}(y)}{dy}\right|$$

Thank You for your Attention!

Thank You for your Attention!

“See you” Next Week