Chapter 5 part 2

Continuous Random Variables

Jingchen (Monika) Hu

Vassar College

MATH 241

Outline

1. Normal distribution
2. Distribution of a function of a continuous random variable

Outline

1. Normal distribution
2. Distribution of a function of a continuous random variable

Normal Distribution

Definition

A continuous random variable X has a Normal distribution distribution with mean μ and variance σ^2 if its pdf is

$$X \sim \mathsf{N}(\mu, \sigma^2) \iff f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \textit{ where } x \in \mathbb{R}$$

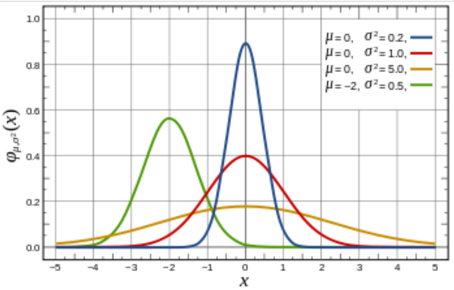
Normal Distribution

Definition

A continuous random variable X has a Normal distribution distribution with mean μ and variance σ^2 if its pdf is

$$X \sim N(\mu, \sigma^2) \iff f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ where } x \in \mathbb{R}$$

Pdf: unimodal and symmetric, bell shaped continuous random variable



The normal pdf is well-defined

(not required)

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Use the fact that normal pdf is well define, calculate the integral

$$\int_{-\infty}^{\infty} e^{-\frac{(x-1)^2}{10}} dx = ?$$

The normal pdf is well-defined

(not required)

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Use the fact that normal pdf is well define, calculate the integral

$$\int_{-\infty}^{\infty} e^{-\frac{(x-1)^2}{10}} dx = ?$$

Suppose we have a random variable $X \sim \mathsf{N}(\mu = 1, \sigma^2 = 5)$, then

$$1 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \cdot 5}}\; e^{-\frac{(x-1)^2}{10}} dx$$

The normal pdf is well-defined

(not required)

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Use the fact that normal pdf is well define, calculate the integral

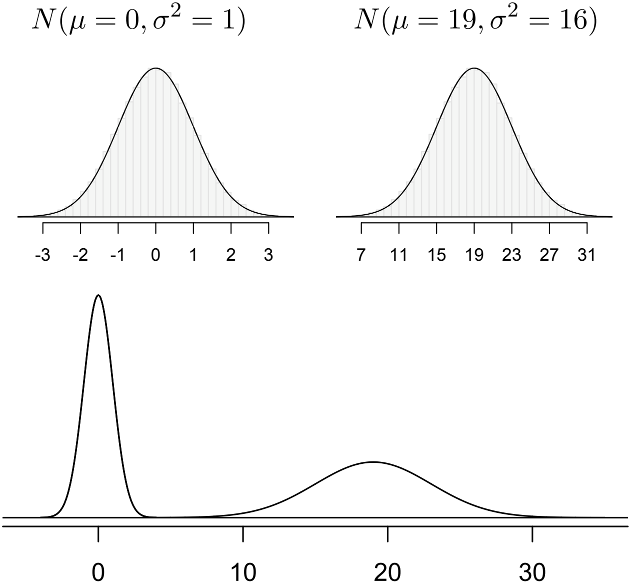
$$\int_{-\infty}^{\infty} e^{-\frac{(x-1)^2}{10}} dx = ?$$

Suppose we have a random variable $X \sim \mathrm{N}(\mu = 1, \sigma^2 = 5)$, then

$$1 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \cdot 5}}\; e^{-\frac{(x-1)^2}{10}} dx$$

$$\Longrightarrow \int_{-\infty}^{\infty} e^{-\frac{(x-1)^2}{10}} dx = \sqrt{10\pi}$$

Normal distributions with different parameters



Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

- $E[X] = \mu$ (required)

Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

- $E[X] = \mu$ (required)

Since $X \sim \mathsf{N}(\mu, \sigma^2) \Longleftrightarrow X = \sigma Z + \mu$ and $Z \sim \mathsf{N}(0, 1)$, it's suffice to show the mean and variance of standard normal are 0 and 1.

Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

- $E[X] = \mu$ (required)

Since $X \sim \mathsf{N}(\mu, \sigma^2) \Longleftrightarrow X = \sigma Z + \mu$ and $Z \sim \mathsf{N}(0, 1)$, it's suffice to show the mean and variance of standard normal are 0 and 1.

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} x f_Z(x) dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \; e^{-\frac{x^2}{2}} dx \\ &= -\frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} = 0 \end{aligned}$$

Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

- $Var(X) = \sigma^2$ (required)

  Since $X \sim \mathsf{N}(\mu, \sigma^2) \iff X = \sigma Z + \mu$ and $Z \sim \mathsf{N}(0, 1)$, it's suffice to show the mean and variance of standard normal are 0 and 1.

Mean and variance of Normal random variable

If $X \sim \mathsf{N}(\mu, \sigma^2)$, and $Z = \frac{X-\mu}{\sigma}$, then Z has a standard normal distribution (more later in Chapter 5.7).

$$Z \sim \mathsf{N}(0, 1)$$

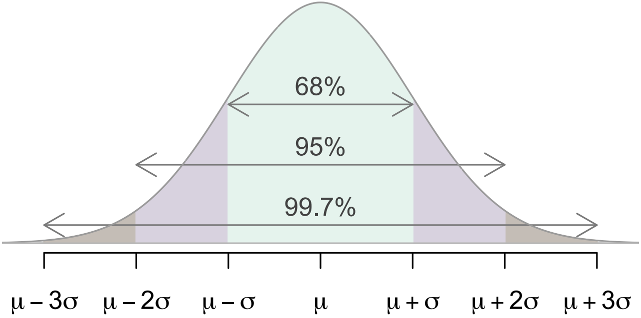
- $Var(X) = \sigma^2$ (required)

  Since $X \sim \mathsf{N}(\mu, \sigma^2) \iff X = \sigma Z + \mu$ and $Z \sim \mathsf{N}(0, 1)$, it's suffice to show the mean and variance of standard normal are 0 and 1.

$$\begin{aligned}
Var[Z] = E[Z^2] - (E[Z])^2 = E[Z^2] &= \int_{-\infty}^{\infty} x^2 f_Z(x)dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \ e^{-\frac{x^2}{2}} dx \ \ (\text{integration by parts} : u = x, dv = xe^{-x^2/2}dx) \\
&= \frac{1}{\sqrt{2\pi}} \left\{ - xe^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right\} \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1
\end{aligned}$$

68-95-99.7 Rule

- A random variable X has a normal distribution,
  - about 68% probability X falls within 1 SD of the mean,
  - about 95% probability X falls within 2 SD of the mean,
  - about 99.7% probability X falls within 3 SD of the mean.
- The probability of X falls 4, 5, or more standard deviations away from the mean is very low.



Normal probability calculation

- We denote $\phi(x)$ and $\Phi(x)$ as pdf and cdf of the standard normal distribution respectively.

Normal probability calculation

- We denote $\phi(x)$ and $\Phi(x)$ as pdf and cdf of the standard normal distribution respectively.
- Probability calculations for X in terms of Z:

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

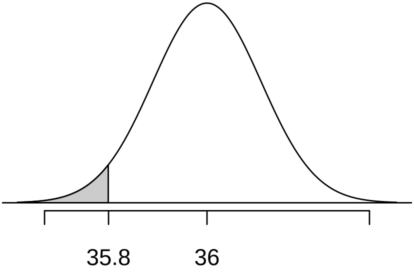
At Heinz ketchup factory the amounts which go into bottles of ketchup are supposed to be normally distributed with mean 36 oz. and standard deviation 0.11 oz. Once every 30 minutes a bottle is selected from the production line, and its contents are noted precisely. If the amount of ketchup in the bottle is below 35.8 oz. or above 36.2 oz., then the bottle will fails the quality control inspection. What's the probability that the amount of ketchup in a randomly selected bottle is less than 35.8 ounces?

At Heinz ketchup factory the amounts which go into bottles of ketchup are supposed to be normally distributed with mean 36 oz. and standard deviation 0.11 oz. Once every 30 minutes a bottle is selected from the production line, and its contents are noted precisely. If the amount of ketchup in the bottle is below 35.8 oz. or above 36.2 oz., then the bottle will fails the quality control inspection. What's the probability that the amount of ketchup in a randomly selected bottle is less than 35.8 ounces?

Let $X =$ amount of ketchup in a bottle: $X \sim N(36, 0.11^2)$

At Heinz ketchup factory the amounts which go into bottles of ketchup are supposed to be normally distributed with mean 36 oz. and standard deviation 0.11 oz. Once every 30 minutes a bottle is selected from the production line, and its contents are noted precisely. If the amount of ketchup in the bottle is below 35.8 oz. or above 36.2 oz., then the bottle will fails the quality control inspection. What's the probability that the amount of ketchup in a randomly selected bottle is less than 35.8 ounces?

Let X = amount of ketchup in a bottle: $X \sim N(36, 0.11^2)$



$$z = \frac{x - \mu}{\sigma} = \frac{35.8 - 36}{0.11} = -1.82$$

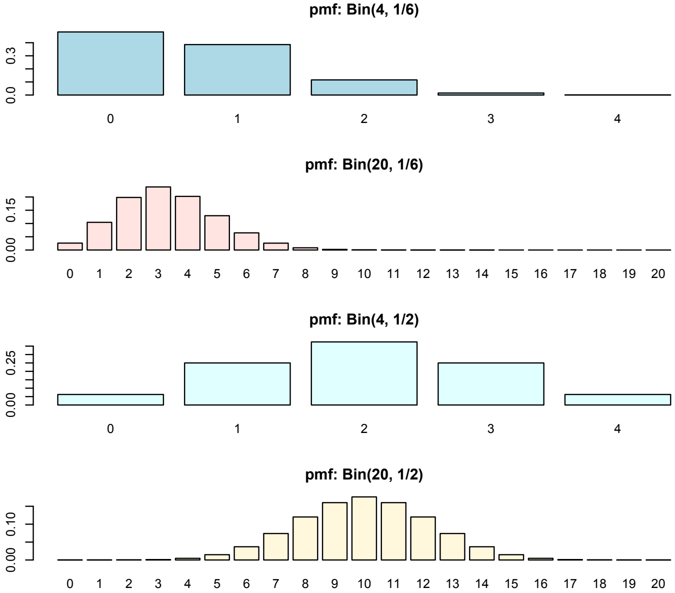
$$P(X < 35.8) = P(Z < -1.82) = 0.0344$$

$$z = \frac{x - \mu}{\sigma} = \frac{35.8 - 36}{0.11} = -1.82$$

$$P(X < 35.8) = P(Z < -1.82) = 0.0344$$

0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	Z
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	*0.0344*	0.0351	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	-1.5

Normal approximation to the Binomial distribution



Normal approximation to the Binomial distribution

Let $X \sim \text{Bin}(n, p)$. When n is large enough, or more specifically, $np(1-p) \geq 10$, the binomial distribution can be approximated by the normal distribution

$$P(X = i) \approx P(i - 0.5 < Y < i + 0.5), \quad Y \sim \text{N}(\mu, \sigma^2)$$

with parameters $\mu = np$ and $\sigma^2 = np(1-p)$.

Normal approximation to the Binomial distribution

Let $X \sim \text{Bin}(n, p)$. When n is large enough, or more specifically, $np(1-p) \geq 10$, the binomial distribution can be approximated by the normal distribution

$$P(X = i) \approx P(i - 0.5 < Y < i + 0.5), \quad Y \sim \text{N}(\mu, \sigma^2)$$

with parameters $\mu = np$ and $\sigma^2 = np(1-p)$.

- Probability calculation (actually, approximation)

$$\begin{aligned} P(a \leq X \leq b) &\approx P(a - 0.5 < Y < b + 0.5) \\ &= P\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}} < \frac{Y - \mu}{\sigma} < \frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) \\ &= \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right) \end{aligned}$$

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

- We are given that $X \sim \text{Bin}(n = 450, p = 0.3)$, and we are asked for the probability

$$P(X > 150) = p(151) + \cdots + p(450) = 1 - p(0) - p(1) - \cdots - p(151)$$

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

- We are given that $X \sim \text{Bin}(n = 450, p = 0.3)$, and we are asked for the probability

$$P(X > 150) = p(151) + \cdots + p(450) = 1 - p(0) - p(1) - \cdots - p(151)$$

- To use normal approximation, first check conditions

$$np(1-p) = 450 \times 0.3 \times 0.7 = 94.5 \geq 10$$

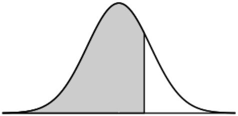
Use Normal approximation with continuity correction,

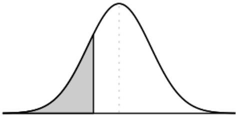
$$\begin{aligned} P(X > 150) &\approx P(X > 150 + 0.5) \\ &= P\left(Z > \frac{150.5 - 450(0.3)}{\sqrt{450(0.3)(0.7)}}\right) \\ &= P(Z > 1.59) \\ &= 1 - 0.9441 = 0.0559 \end{aligned}$$

Recap: Normal distribution $X \sim \mathrm{N}(\mu, \sigma^2)$

- Mean μ, variance σ^2
- Symmetry

$$f(\mu - x) = f(\mu + x), F(\mu - x) = 1 - F(\mu + x)$$

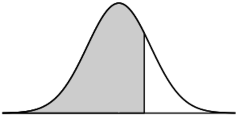




Recap: Normal distribution $X \sim \mathsf{N}(\mu, \sigma^2)$

- Mean μ, variance σ^2
- Symmetry

$$f(\mu - x) = f(\mu + x), F(\mu - x) = 1 - F(\mu + x)$$

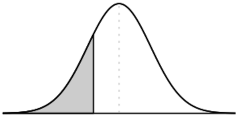


- Standard normal distribution $\mu = 0, \sigma^2 = 1.$

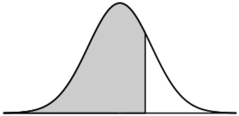
$$X \sim \mathsf{N}(\mu, \sigma^2) \Longleftrightarrow Z = \frac{X - \mu}{\sigma} \sim \mathsf{N}(0, 1)$$

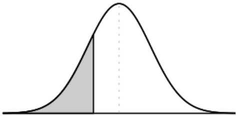
- Find probability using $\Phi(\cdot)$ table

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$



Recap: Normal distribution $X \sim \mathsf{N}(\mu, \sigma^2)$





- Mean μ, variance σ^2
- Symmetry

$$f(\mu - x) = f(\mu + x), F(\mu - x) = 1 - F(\mu + x)$$

- Standard normal distribution $\mu = 0, \sigma^2 = 1.$

$$X \sim \mathsf{N}(\mu, \sigma^2) \Longleftrightarrow Z = \frac{X - \mu}{\sigma} \sim \mathsf{N}(0, 1)$$

- Find probability using $\Phi(\cdot)$ table

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- Normal approximation to Binomial $X \sim \mathsf{Bin}(n, p)$

$$Y \sim \mathsf{N}(\mu = np, \sigma^2 = np(1 - p))$$

$$P(X = i) \approx P(i - 0.5 < Y < i + 0.5)$$

Review: continuous distributions

Name	Range	pdf $f(x)$	mean	variance
Unif(α, β)	$[\alpha, \beta]$	$\frac{1}{\beta-\alpha}$	$\frac{\alpha+\beta}{2}$	$\frac{(\beta-\alpha)^2}{12}$
N(μ, σ^2)	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$	μ	σ^2

Outline

1. Normal distribution

2. Distribution of a function of a continuous random variable

A (important!) theorem on finding pdf of $g(X)$

Suppose X is a continuous random variable with pdf $f_X(x)$. If a function $g(x)$ is

1. monotonic (increasing or decreasing), and
2. differentiable (and thus continuous),

A (important!) theorem on finding pdf of $g(X)$

Suppose X is a continuous random variable with pdf $f_X(x)$. If a function $g(x)$ is

1. monotonic (increasing or decreasing), and
2. differentiable (and thus continuous),

then the random variable defined by $Y = g(X)$ has pdf

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

A (important!) theorem on finding pdf of $g(X)$

Suppose X is a continuous random variable with pdf $f_X(x)$. If a function $g(x)$ is

1. monotonic (increasing or decreasing), and
2. differentiable (and thus continuous),

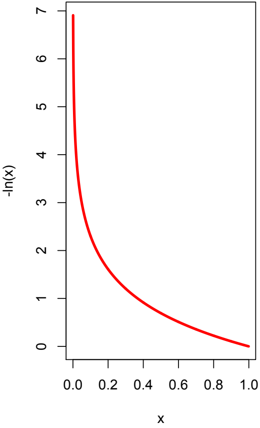
then the random variable defined by $Y = g(X)$ has pdf

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

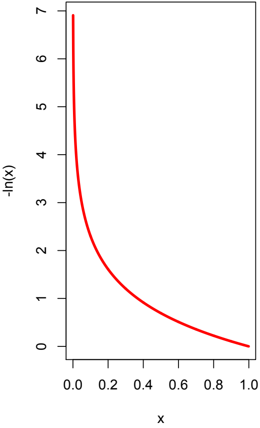
Or more rigorously,

$$f_Y(y) = \begin{cases} f_X\left[g^{-1}(y)\right] \cdot \left| \frac{d}{dy} g^{-1}(y) \right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

Example: let $X \sim \text{Unif}(0,1)$, what distribution does $Y = -\ln(X)$ have?



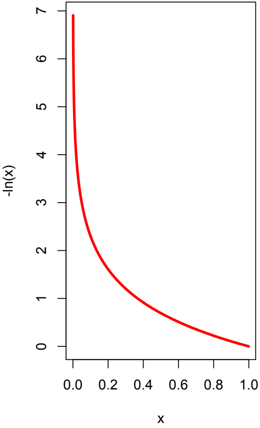
Example: let $X \sim \text{Unif}(0, 1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic
2. differentiable

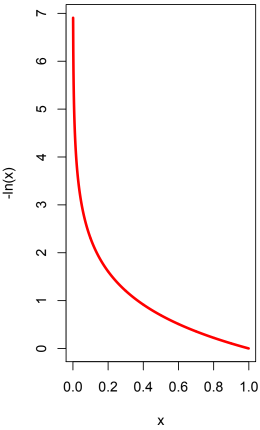
Example: let $X \sim \text{Unif}(0, 1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic ✓
2. differentiable

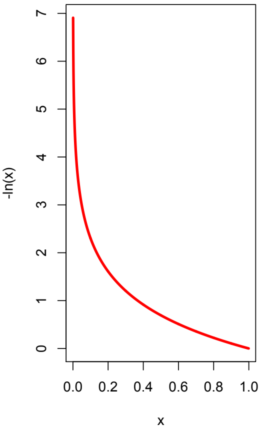
Example: let $X \sim \text{Unif}(0, 1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic ✓
2. differentiable ✓

Example: let $X \sim$ Unif$(0, 1)$, what distribution does $Y = -\ln(X)$ have?



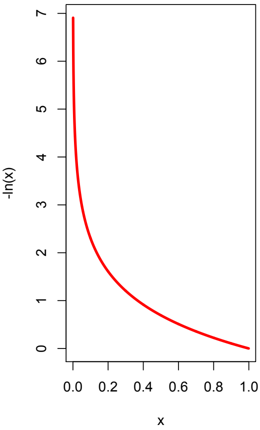
In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic ✓
2. differentiable ✓

Inverse function $g^{-1}(y)$:

$$y = -\ln(x) \iff x = e^{-y}$$

Example: let $X \sim \text{Unif}(0, 1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

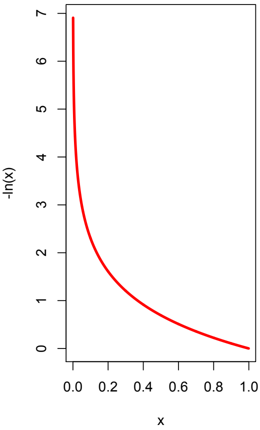
1. monotonic ✓
2. differentiable ✓

Inverse function $g^{-1}(y)$:

$$y = -\ln(x) \Longleftrightarrow x = e^{-y}$$

$$\frac{dx}{dy} = -e^{-y}$$

Example: let $X \sim \text{Unif}(0,1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic ✓
2. differentiable ✓

Inverse function $g^{-1}(y)$:

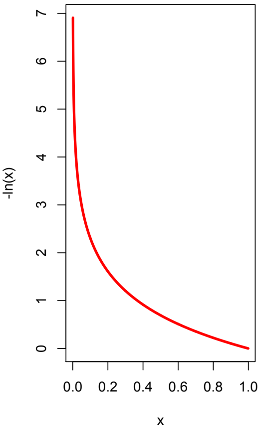
$$y = -\ln(x) \iff x = e^{-y}$$

$$\frac{dx}{dy} = -e^{-y}$$

Range of Y: $y \in [0, \infty)$.

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = 1 \cdot \left|-e^{-y}\right| = e^{-y}$$

Example: let $X \sim \text{Unif}(0,1)$, what distribution does $Y = -\ln(X)$ have?



In order to use the previous theorem, need to check the function $g(x) = -\ln(x)$

1. monotonic ✓
2. differentiable ✓

Inverse function $g^{-1}(y)$:

$$y = -\ln(x) \iff x = e^{-y}$$

$$\frac{dx}{dy} = -e^{-y}$$

Range of Y: $y \in [0, \infty)$.

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right| = 1 \cdot \left| -e^{-y} \right| = e^{-y}$$

Therefor, Y has an exponential distribution,

$$Y \sim \text{Exp}(1)$$

Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|$?

a. Yes
b. No

Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$?

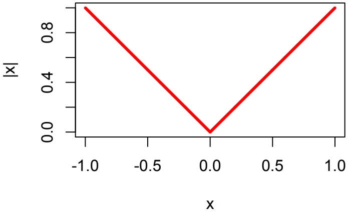
(a) Yes
(b) *No*

Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$?

(a) Yes

(b) *No*

$g(x) = |x|$ is not monotonic; also not differentiable at 0.

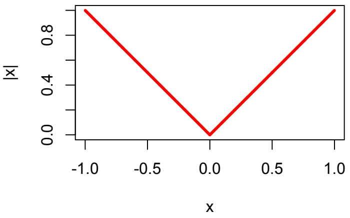


Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$?

(a) Yes
(b) *No*

$g(x) = |x|$ is not monotonic; also not differentiable at 0.

First find cdf $F_Y(y)$, then find $f_Y(y)$.

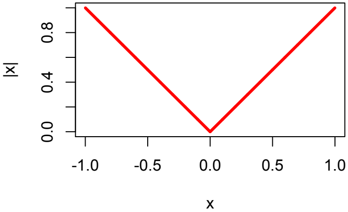


Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$?

(a) Yes
(b) *No*

$g(x) = |x|$ is not monotonic; also not differentiable at 0.

First find cdf $F_Y(y)$, then find $f_Y(y)$.
For any $y \geq 0$,



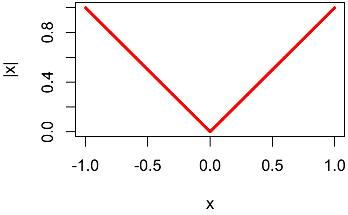
$$
\begin{aligned}
F_Y(y) &= P(Y < y) \\
&= P(|X| < y) \\
&= P(-y < X < y) \\
&= F_X(y) - F_X(-y)
\end{aligned}
$$

Let X have pdf $f_X(x)$, where $-\infty < x < \infty$. We want to find the pdf for $Y = |X|$. Can we use the formula $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$?

(a) Yes
(b) *No*

$g(x) = |x|$ is not monotonic; also not differentiable at 0.

First find cdf $F_Y(y)$, then find $f_Y(y)$.
For any $y \geq 0$,



$$
\begin{aligned}
F_Y(y) &= P(Y < y) \\
&= P(|X| < y) \\
&= P(-y < X < y) \\
&= F_X(y) - F_X(-y)
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y) &= \frac{d}{dy}[F_X(y) - F_X(-y)] \\
&= f_X(y) + f_X(-y)
\end{aligned}
$$