

Continuous Random Variables

Dr. Edmund Lam

Department of Electrical and Electronic Engineering
The University of Hong Kong

ELEC2844: Probabilistic Systems Analysis

(Second Semester, 2020–21)

<https://www.eee.hku.hk/~elec2844>

Random variables

Continuous random variable X :

- Normalization:

$$\mathbb{P}(-\infty < X < \infty) = \int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (3)$$

- For an interval $[x, x + \delta]$:

$$\mathbb{P}([x, x + \delta]) = \int_x^{x+\delta} f_X(t) dt \approx f_X(x) \cdot \delta$$

So $f_X(x)$ is the “probability mass per unit length” near x

- Since $f_X(x)$ is not the probability of any event, it does not have to be ≤ 1 .

Random variables

Continuous random variable X :

- has a nonnegative function

$$f_X(x) \geq 0 \quad (1)$$

for all x , called **probability density function (PDF)**, such that
 $\mathbb{P}(X \in B) = \int_B f_X(x) dx$ for every subset B of the real line

- For B to refer to a range $[a, b]$,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx \quad (2)$$

- Integral is zero when the range contains only a single number, so

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X \leq b)$$

Example

Is $f_X(x)$ a PDF?

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{x}} & 0 < x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

ANS: Even though $f_X(x)$ becomes infinitely large as $x \rightarrow 0$, we still have $f_X(x) \geq 0$ for all x , and

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_0^1 \frac{1}{2\sqrt{x}} dx = \sqrt{x} \Big|_0^1 = 1$$

So this is a PDF.

- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 Derived distributions
- 5 Joint PDF, conditioning, independence
- 6 Continuous Bayes' rule
- 7 (Advanced Topic) Moment generating function

Expectation and variance

The formulas for continuous r.v. are similar to those for discrete:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx \quad (5)$$

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (6)$$

$$\text{var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 f_X(x) dx \quad (7)$$

$$\text{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \geq 0 \quad (8)$$

$$\mathbb{E}(Y) = a\mathbb{E}(X) + b \quad (9)$$

$$\text{var}(Y) = a^2 \text{var}(X) \quad (10)$$

where $Y = aX + b$

Expectation

The **expectation** or **expected value** or **mean** of a continuous r.v. X is

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (4)$$

Similar to the discrete case, it is a very useful number as a “representative” value of X , and is also viewed as the “center of gravity.”

Note: There can be problematic cases with integrals: even if $f_X(x)$ is symmetric about 0, the expected value is in fact not defined, e.g. $f_X(x) = c/(1+x^2)$

- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 Derived distributions
- 5 Joint PDF, conditioning, independence
- 6 Continuous Bayes' rule
- 7 (Advanced Topic) Moment generating function

Continuous random variable

It is useful to know the mean and variance of some common distributions. A summary:

	mean: $\mathbb{E}(X)$	variance: $\text{var}(X)$
uniform: $X \sim U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponential: $X \sim \text{Exp}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
normal: $X \sim \mathcal{N}(\mu, \sigma^2)$	μ	σ^2

(1) Uniform random variable

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{2} x^2 \Big|_a^b = \frac{a+b}{2}$$

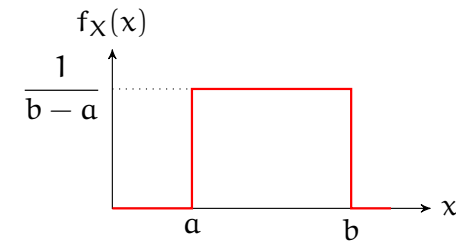
$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{3} x^3 \Big|_a^b = \frac{a^2 + ab + b^2}{3}$$

$$\text{var}(X) = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

(1) Uniform random variable

Continuous uniform random variable over a range $[a, b]$ has PDF

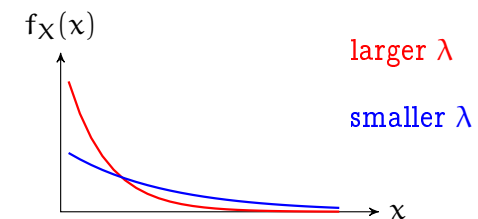
$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$



(2) Exponential random variable

Exponential random variable with parameter $\lambda > 0$ has PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$



(2) Exponential random variable

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E}(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = (-x e^{-\lambda x}) \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = 0 - \frac{e^{-\lambda x}}{\lambda} \Big|_0^{\infty} = \frac{1}{\lambda}$$

$$\mathbb{E}(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = (-x^2 e^{-\lambda x}) \Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx = \frac{2}{\lambda} \mathbb{E}(X) = \frac{2}{\lambda^2}$$

$$\text{var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

(3) Normal random variable

Show that the normal distribution given by Eq. (13) has unit area.

Proof: Let $I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$. The trick is to go two-dimensional by computing I^2 .

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2/2} dy \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

Then, we convert to polar coordinates:

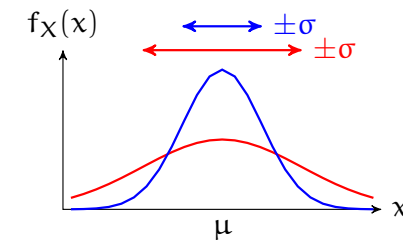
$$\begin{aligned} I^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta = 2\pi \int_0^{\infty} e^{-r^2/2} r dr \\ &= 2\pi \int_0^{\infty} e^{-u} du = 2\pi (-e^{-u}) \Big|_0^{\infty} = 2\pi \\ I &= \sqrt{2\pi} \end{aligned}$$

(3) Normal random variable

Gaussian or normal r.v. with parameters $(\mu, \sigma > 0)$ has PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad (13)$$

This is probably the most common distribution of all (with good scientific reason)!



(3) Normal random variable

Thus, we have shown that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-s^2/2} ds = 1$$

Then, we make the change of variable $s = \frac{x-\mu}{\sigma}$:

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

since $ds = dx/\sigma$.

(3) Normal random variable

Normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ has **mean** μ and **variance** σ^2 .

Proof: For the mean, we will appeal to the fact that it is symmetric around μ and that the integral is defined. For variance, making use of a change of variable $y = (x - \mu)/\sigma$,

$$\begin{aligned}\text{var}(X) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left[(-ye^{-y^2/2}) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-y^2/2} dy \right] \\ &= \frac{\sigma^2}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) = \sigma^2\end{aligned}$$

1 Expectation, mean, variance

2 Common continuous random variables

3 Cumulative density function

4 Derived distributions

5 Joint PDF, conditioning, independence

6 Continuous Bayes' rule

7 (Advanced Topic) Moment generating function

(3) Normal random variable

If $X \sim \mathcal{N}(\mu, \sigma^2)$, and a linear transformation $Y = aX + b$ where $a \neq 0$ and b are scalars,

$$\begin{cases} X \sim \mathcal{N}(\mu, \sigma^2) \\ Y = aX + b \end{cases} \implies Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2) \quad (14)$$

Corollary:

If $Y = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$, then $Y \sim \mathcal{N}(0, 1)$, called **standard normal**.

To see why, we'll look at cumulative density function and its use in deriving transforms of random variables first.

Definition

The **cumulative density function (CDF)** of a random variable X is denoted $F_X(x)$. This can be defined for both discrete and continuous r.v.:

$$F_X(x) = \mathbb{P}(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(k) & X \text{ is discrete} \\ \int_{-\infty}^x f_X(t) dt & X \text{ is continuous} \end{cases} \quad (15)$$

It is more common to use it for the continuous case.

Interpretation: CDF is the accumulated probability up to x .

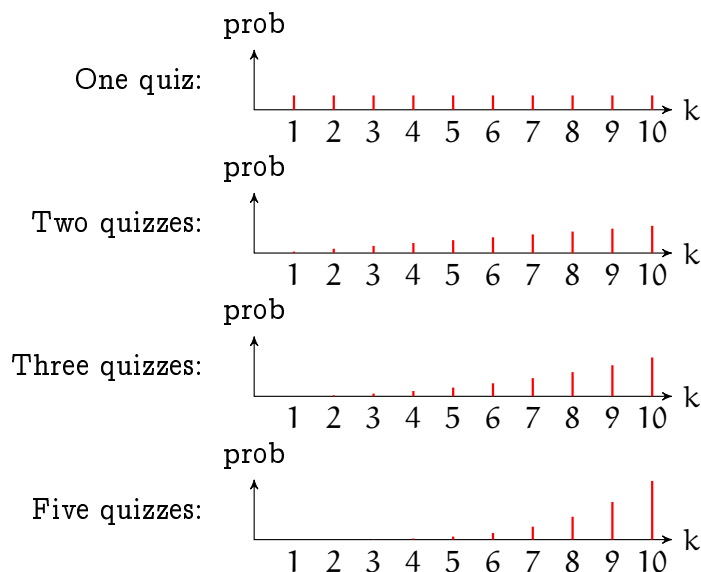
Definition

We can also obtain PMF or PDF from CDF:

$$\begin{cases} p_X(k) = F_X(k) - F_X(k-1) & \text{discrete} \\ f_X(x) = \frac{d}{dx} F_X(x) & \text{continuous} \end{cases} \quad (16)$$

This is particularly useful when we define new r.v. through CDF, and then differentiate (for continuous r.v.) to obtain PDF.

Example



Example

Assume you take a quiz and the score is (discrete) uniform between 1 and 10. You are allowed to take the test three times and count only their maximum. What is the final PMF?

ANS: Let the quiz scores be X_1, X_2, X_3 , so that $X = \max\{X_1, X_2, X_3\}$.

$$\begin{aligned} F_X(k) &= \mathbb{P}(X \leq k) = \mathbb{P}(X_1 \leq k, X_2 \leq k, X_3 \leq k) \\ &= \mathbb{P}(X_1 \leq k) \mathbb{P}(X_2 \leq k) \mathbb{P}(X_3 \leq k) = \left(\frac{k}{10}\right)^3 \end{aligned}$$

since the quiz score for each test is independent of each other.

$$\begin{aligned} p_X(k) &= F_X(k) - F_X(k-1) \\ &= \left(\frac{k}{10}\right)^3 - \left(\frac{k-1}{10}\right)^3 \quad k = 1, \dots, 10 \end{aligned}$$

Geometric CDF

Remember $X \sim \text{Geo}(p)$ has PMF

$$p_X(k) = (1-p)^{k-1}p \quad k = 1, 2, \dots$$

Its CDF is

$$F_X(n) = \sum_{k=1}^n (1-p)^{k-1}p = \frac{1 - (1-p)^n}{1 - (1-p)}p = 1 - (1-p)^n \quad (17)$$

for $n = 1, 2, \dots$

Exponential CDF

$X \sim \text{Exp}(\lambda)$ has PDF

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0$$

Its CDF is

$$F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x} \quad (18)$$

for $x > 0$.

Normal CDF

For $X \sim \mathcal{N}(0, 1)$, the CDF is

$$F_X(x) = \Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (20)$$

We cannot analytically integrate the function, but we can compute it as a function $\Phi(x)$ beforehand.

Then, for $X \sim \mathcal{N}(\mu, \sigma^2)$, we can “standardize” by defining $Y = \frac{X - \mu}{\sigma}$ such that $\mathbb{E}(Y) = 0$ and $\text{var}(Y) = 1$, and

$$\mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \mathbb{P}\left(Y \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right) \quad (21)$$

Geometric and exponential CDF

Geometric and exponential CDFs are closely related. By defining $\delta = -\ln(1 - p)/\lambda$, we have

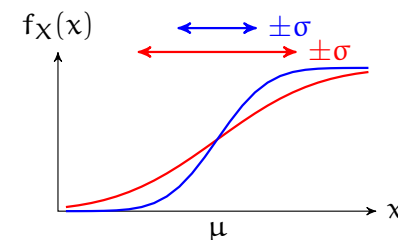
$$e^{-\lambda \delta} = 1 - p$$

and whenever $x = n\delta$,

$$F_{\text{exp}}(n\delta) = 1 - e^{-\lambda n\delta} = 1 - (1 - p)^n = F_{\text{geo}}(n) \quad (19)$$

Interpretation: Suppose we toss very quickly (every δ seconds, where $\delta \ll 1$) a biased coin with a very small probability of heads (equal to $p = 1 - e^{-\lambda \delta}$). The first time to obtain head, which is a geometric r.v. with parameter p , is a close approximation to an exponential r.v. with parameter λ .

(3) Normal CDF



Example

A binary signal $\{+1, -1\}$ is transmitted through a channel corrupted with additive Gaussian noise $N \sim \mathcal{N}(0, \sigma^2)$. The receiver decides the signal is $+1$ if the received signal is ≥ 0 , and -1 if the received signal is < 0 . What is the probability of error?

ANS: When the signal is -1 , error occurs if $N \geq 1$.

$$\mathbb{P}(N \geq 1) = 1 - \mathbb{P}(N < 1) = 1 - \Phi\left(\frac{1}{\sigma}\right)$$

By symmetry, we have the same probability of error when signal is $+1$, and error occurs if $N < -1$.

As example, when $\sigma = 1$, $\Phi(1) \approx 0.84$, and probability of error is ≈ 0.16 .

Procedure

Given: PDF of X , and $Y = g(X)$, find the PDF of Y .

A two-step procedure:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y) = \int_{\{x \mid g(x) \leq y\}} f_X(x) dx \quad (22)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) \quad (23)$$

- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 **Derived distributions**
- 5 Joint PDF, conditioning, independence
- 6 Continuous Bayes' rule
- 7 (Advanced Topic) Moment generating function

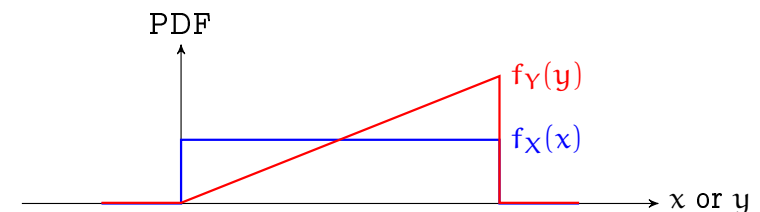
Example

$X \sim U(0, 1)$ and $Y = \sqrt{X}$. Find PDF of Y .

ANS: Follow the procedure to find $F_Y(y)$ and $f_Y(y)$.

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\sqrt{X} \leq y) = \mathbb{P}(X \leq y^2) = y^2$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} (y^2) = 2y \quad 0 \leq y \leq 1$$



Example

$X \sim U(5, 10)$ and $Y = \frac{1}{X}$. Find PDF of Y .

ANS: Procedure is similar, but we need $F_X(x)$ first.

$$F_X(x) = \begin{cases} 0 & x \leq 5 \\ \frac{x-5}{5} & 5 < x \leq 10 \\ 1 & 10 < x \end{cases}$$

Then, $F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(\frac{1}{X} \leq y\right) = \mathbb{P}\left(\frac{1}{y} \leq X\right) = 1 - F_X\left(\frac{1}{y}\right)$.

$$F_Y(y) = \begin{cases} 0 & y \leq \frac{1}{10} \\ 1 - \left(\frac{1}{y} - 5\right)/5 = 2 - \frac{1}{5y} & \frac{1}{10} < y \leq \frac{1}{5} \\ 1 & \frac{1}{5} < y \end{cases}$$

Example

X has known PDF, and $Y = X^2$. Find PDF of Y .

ANS: For $y \geq 0$,

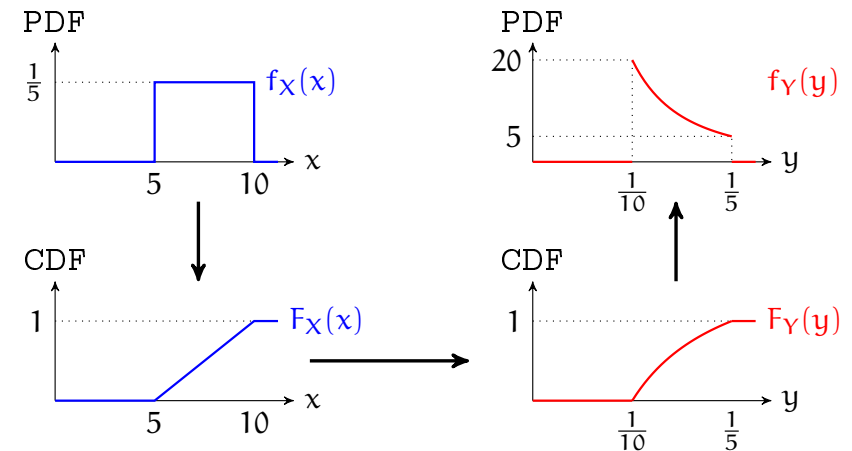
$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

Differentiating,

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}) \quad y \geq 0$$

Example

Differentiating, $f_Y(y) = \begin{cases} 0 & y \leq \frac{1}{10} \text{ or } \frac{1}{5} < y \\ \frac{1}{5y^2} & \frac{1}{10} < y \leq \frac{1}{5} \end{cases}$



Special case: Y is a linear function of X

Given: PDF of X , and $Y = aX + b$, find the PDF of Y .

Case 1: $a > 0$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(aX + b \leq y) = \mathbb{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Case 2: $a < 0$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(aX + b \leq y) = \mathbb{P}\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -\frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Special case: Y is a linear function of X

Given: PDF of X , and $Y = aX + b$,

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \quad (24)$$

Example: $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{y-b}{a} - \mu\right)^2 / 2\sigma^2} \\ &= \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}} \end{aligned}$$

i.e., $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Definition

Consider two continuous r.v. X and Y , with **joint PDF**

$$\mathbb{P}((X, Y) \in B) = \iint_{(x,y) \in B} f_{X,Y}(x, y) \, dx \, dy \quad (25)$$

for every subset B of the 2D plane.

When $B = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$,

$$\mathbb{P}(a \leq x \leq b, c \leq y \leq d) = \int_c^d \int_a^b f_{X,Y}(x, y) \, dx \, dy \quad (26)$$

- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 Derived distributions
- 5 Joint PDF, conditioning, independence**
- 6 Continuous Bayes' rule
- 7 (Advanced Topic) Moment generating function

Definition

We can calculate the **marginal PDFs** of X and Y using

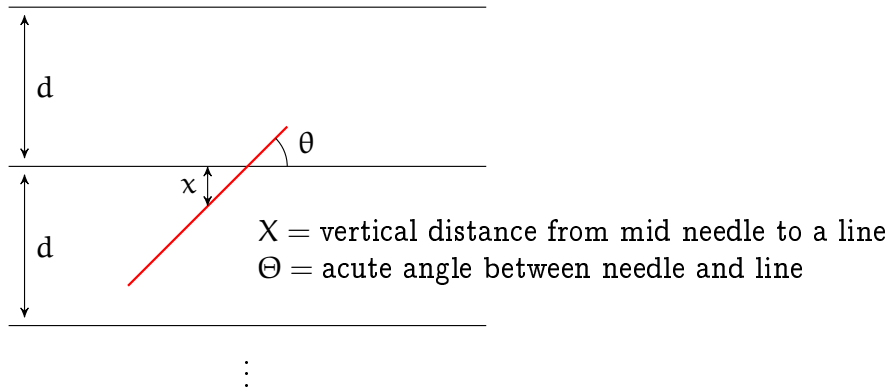
$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dy \quad (27)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dx \quad (28)$$

and, for normalization, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dx \, dy = 1$

Example

Buffon's Needle. We have many parallel lines at distances d apart. Assume we throw a needle of length l , where $l < d$, on the surface at random. What is the probability that the needle will intersect one of the lines?



Example

$$\begin{aligned}\mathbb{P}\left(X \leq \frac{l}{2} \sin \Theta\right) &= \int_0^{\pi/2} \int_0^{(l/2) \sin \theta} \frac{4}{\pi d} dx d\theta \\ &= \frac{4}{\pi d} \int_0^{\pi/2} \frac{l}{2} \sin \theta d\theta \\ &= \frac{2l}{\pi d} (-\cos \theta) \Big|_0^{\pi/2} \\ &= \frac{2l}{\pi d}\end{aligned}$$

This can be used as a **Monte Carlo method** to find π : if we drop n needles and h of them cross lines, then

$$\pi \approx \frac{2l}{d} \cdot \frac{n}{h}$$

Example

At random: assume a uniform joint PDF over the rectangular set

$$S = \left\{ (x, \theta) \mid 0 \leq x \leq \frac{d}{2}, 0 \leq \theta \leq \frac{\pi}{2} \right\}$$

This rectangle has an area $\pi d/4$, and X and Θ are independent, so

$$f_{X,\Theta}(x, \theta) = \begin{cases} \frac{4}{\pi d} & (x, \theta) \in S \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, $\frac{l}{2} \sin \Theta$ is the *vertical* distance between the middle of the needle to the end points of the needle; the needle will intersect one of the lines if and only if

$$X \leq \frac{l}{2} \sin \Theta$$

Conditioning on an event

We can have **conditional PDF** of a continuous r.v. X given event A

$$\mathbb{P}(X \in B | A) = \int_B f_{X|A}(x) dx \quad (29)$$

with $\int_{-\infty}^{\infty} f_{X|A}(x) dx = 1$. Specifically, if the event is of the form $\{X \in A\}$ with $\mathbb{P}(X \in A) > 0$,

$$\mathbb{P}(X \in B | X \in A) = \frac{\mathbb{P}(X \in B, X \in A)}{\mathbb{P}(X \in A)} = \frac{\int_{A \cap B} f_X(x) dx}{\mathbb{P}(X \in A)}$$

$$f_{X|\{X \in A\}}(x) = \begin{cases} \frac{f_X(x)}{\mathbb{P}(X \in A)} & x \in A \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

Interpretation: Integrate all cases of $X = x$ AND belonging to A , then normalize with $\mathbb{P}(X \in A)$.

Example

A certain device has a lifetime following an exponential r.v. with parameter λ . Suppose now we know it has remained in operation for time t ; and let X be the *additional* time until it fails. What is the conditional PDF of X , given it has already lasted for time t ?

ANS: Define event $A = \{T > t\}$. Let us look at $\mathbb{P}(X > x | A)$.

$$\begin{aligned}\mathbb{P}(X > x | A) &= \mathbb{P}(T > t + x | T > t) = \frac{\mathbb{P}(T > t + x \text{ and } T > t)}{\mathbb{P}(T > t)} \\ &= \frac{\mathbb{P}(T > t + x)}{\mathbb{P}(T > t)} \\ &= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x}\end{aligned}$$

So $\mathbb{P}(X < x | A) = 1 - e^{-\lambda x}$, which means $X \sim \text{Exp}(\lambda)$.

Interpretation: Exponential r.v. is memoryless.

Example

A dart lands uniformly random on a circular target of radius r . Let x and y be the Cartesian coordinates. Find the marginal PDF $f_Y(y)$ and conditional PDF $f_{X|Y}(x|y)$.

ANS: The joint PDF is $f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi r^2} & x^2 + y^2 \leq r^2 \\ 0 & \text{otherwise.} \end{cases}$

The marginal PDF is

$$\begin{aligned}f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \frac{1}{\pi r^2} \int_{x^2 + y^2 \leq r^2} dx \\ &= \frac{1}{\pi r^2} \int_{-\sqrt{r^2 - y^2}}^{\sqrt{r^2 - y^2}} dx \\ &= \frac{2}{\pi r^2} \sqrt{r^2 - y^2} \quad \text{if } |y| \leq r\end{aligned}$$

This is NOT uniform!

Conditioning on a r.v.

The **conditional PDF** of a r.v. X given knowledge of another r.v. Y

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (31)$$

$$\text{similarly, } \int_{-\infty}^{\infty} f_{X|Y}(x|y) dx = 1.$$

Example

The conditional PDF is

$$\begin{aligned}f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{\frac{1}{\pi r^2}}{\frac{2}{\pi r^2} \sqrt{r^2 - y^2}} \\ &= \frac{1}{2\sqrt{r^2 - y^2}} \quad \text{if } x^2 + y^2 \leq r^2\end{aligned}$$

Given a fixed y , this is uniform!

Conditional expectation

Let X and Y be two continuous r.v., and A an event. We can compute the conditional expectations:

$$\mathbb{E}(X|A) = \int_{-\infty}^{\infty} x f_{X|A}(x) dx \quad (32)$$

$$\mathbb{E}(g(X)|A) = \int_{-\infty}^{\infty} g(x) f_{X|A}(x) dx \quad (33)$$

$$\mathbb{E}(X|Y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \quad (34)$$

$$\mathbb{E}(g(X)|Y) = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx \quad (35)$$

Conditional expectation

Extensions to functions of several r.v.:

$$\mathbb{E}(g(X, Y)|Y) = \int_{-\infty}^{\infty} g(x, y) f_{X|Y}(x|y) dx \quad (38)$$

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{\infty} \mathbb{E}(g(X, Y)|Y) f_Y(y) dy \quad (39)$$

Conditional expectation

Assume $A_1, \dots, A_n$ are disjoint events partitioning the sample space, with $\mathbb{P}(A_i) > 0$ for all i .

$$\mathbb{E}(X) = \sum_{i=1}^n \mathbb{P}(A_i) \mathbb{E}(X|A_i) \quad (36)$$

unconditional average = averaging the conditional averages

Also when $A_i = \{Y = y\}$,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} \mathbb{E}(X|Y = y) f_Y(y) dy \quad (37)$$

Independence

Two r.v. X and Y are independent of each other,

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \quad (40)$$

Since $f_{X,Y}(x, y) = f_{X|Y}(x|y) f_Y(y)$, for all y with $f_Y(y) > 0$,

$$f_{X|Y}(x|y) = f_X(x) \quad (41)$$

Can also interchange X and Y . Furthermore,

$$\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y) \quad (42)$$

$$\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X)) \mathbb{E}(h(Y)) \quad (43)$$

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) \quad (44)$$

- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 Derived distributions
- 5 Joint PDF, conditioning, independence
- 6 Continuous Bayes' rule**
- 7 (Advanced Topic) Moment generating function

Example

Suppose we have electronic devices with exponential lifetime $Y \sim \text{Exp}(\lambda)$, but λ itself is a random variable, modeled by uniform distribution $\Lambda \sim U(1, 3)$. Given we observe the lifetime of a particular device to be $Y = y$, what do we know about λ ?

ANS: We want to find $f_{\Lambda|Y}(\lambda|y)$ through Bayes' rule.

$$f_{\Lambda}(\lambda) = \frac{1}{3-1} = \frac{1}{2} \quad 1 \leq \lambda \leq 3$$

$$f_{Y|\Lambda}(y|\lambda) = \lambda e^{-\lambda y}$$

$$\int_{-\infty}^{\infty} f_{\Lambda}(t) f_{Y|\Lambda}(y|t) dt = \int_1^3 \frac{1}{2} t e^{-ty} dt$$

Definition

Unobserved phenomenon: r.v. X with $f_X(x)$

Measurement: Y , related to X through $f_{Y|X}(y|x)$

From the observation $Y = y$, what do we know about X ?

$\Rightarrow$ we want to know $f_{X|Y}(x|y)$

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{f_Y(y)} \quad (45)$$

This is the continuous form of Bayes' rule. Alternatively,

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(t) f_{Y|X}(y|t) dt} \quad (46)$$

Example

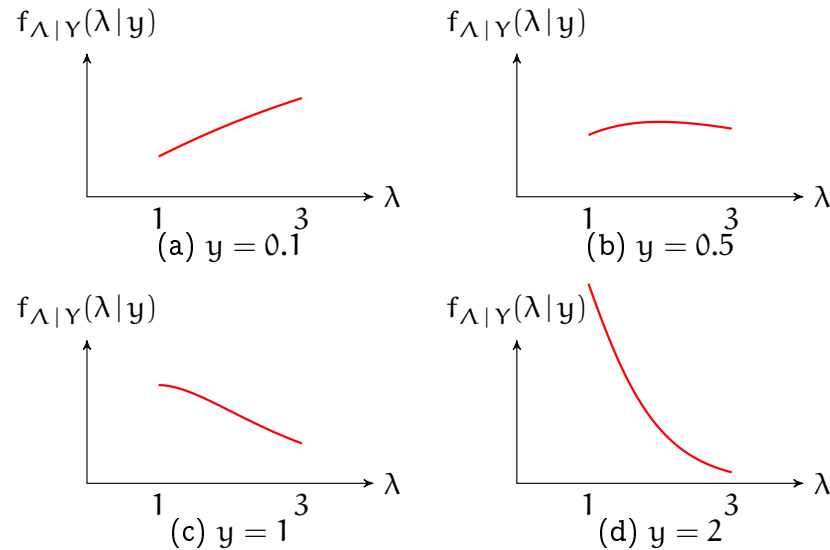
Integration by parts:

$$\int_a^b v \frac{du}{dt} dt = uv \Big|_a^b - \int_a^b u \frac{dv}{dt} dt$$

Let $v = t$ and $u = -\frac{1}{y} e^{-yt}$, $a = 1$ and $b = 3$,

$$\begin{aligned} \frac{1}{2} \int_1^3 t e^{-yt} dt &= -\frac{t}{2y} e^{-yt} \Big|_1^3 - \int_1^3 -\frac{1}{2y} e^{-yt} dt \\ &= \frac{1}{2} \left(-\frac{3}{y} e^{-3y} + \frac{1}{y} e^{-y} \right) - \frac{1}{2y^2} e^{-yt} \Big|_1^3 \\ &= \frac{1}{2y} (e^{-y} - 3e^{-3y}) + \frac{1}{2y^2} (e^{-y} - e^{-3y}) \\ f_{\Lambda|Y}(\lambda|y) &= \frac{\lambda e^{-\lambda y}}{\frac{1}{y} (e^{-y} - 3e^{-3y}) + \frac{1}{y^2} (e^{-y} - e^{-3y})} \quad 1 \leq \lambda \leq 3 \end{aligned}$$

Example



Special case 1: Unobserved phenomenon is discrete

Case 1a: A is “binary”, so $f_Y(y) = \mathbb{P}(A) f_{Y|A}(y) + \mathbb{P}(A^c) f_{Y|A^c}(y)$

$$\mathbb{P}(A|Y=y) = \frac{\mathbb{P}(A) f_{Y|A}(y)}{\mathbb{P}(A) f_{Y|A}(y) + \mathbb{P}(A^c) f_{Y|A^c}(y)} \quad (48)$$

Case 1b: A is of the form $\{N = n\}$, where $n = 1, \dots, K$, with $p_N(n)$

$$\mathbb{P}(N = n|Y = y) = \frac{p_N(n) f_{Y|N}(y|n)}{\sum_{i=1}^K p_N(i) f_{Y|N}(y|i)} \quad (49)$$

Special case 1: Unobserved phenomenon is discrete

Unobserved phenomenon: event A with $\mathbb{P}(A)$

Measurement: Y , related to A through $f_{Y|A}(y)$

From the observation $Y = y$, what do we know about A ?

$\Rightarrow$ we want to know $\mathbb{P}(A|Y=y)$

$$\mathbb{P}(A|Y=y) = \frac{\mathbb{P}(A) f_{Y|A}(y)}{f_Y(y)} \quad (47)$$

Example

A binary signal S with $\mathbb{P}(S = 1) = p$ and $\mathbb{P}(S = -1) = 1 - p$ is transmitted, which is corrupted with independent Gaussian noise $N \sim \mathcal{N}(0, 1)$, and the received signal is $Y = N + S$. What is the probability of $S = 1$ given an observation $Y = y$?

ANS: Let $A = \{S = 1\}$.

$$f_{Y|A}(y) \sim \mathcal{N}(1, 1)$$

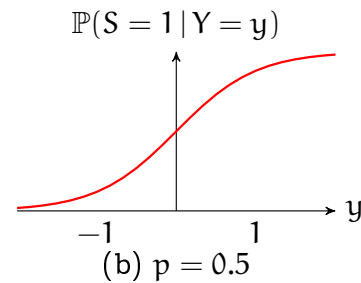
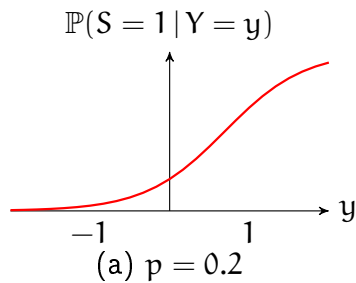
$$f_{Y|A^c}(y) \sim \mathcal{N}(-1, 1)$$

$$\begin{aligned} \mathbb{P}(A|Y=y) &= \frac{p \cdot \frac{1}{\sqrt{2\pi}} e^{-(y-1)^2/2}}{p \cdot \frac{1}{\sqrt{2\pi}} e^{-(y-1)^2/2} + (1-p) \cdot \frac{1}{\sqrt{2\pi}} e^{-(y+1)^2/2}} \\ &= \frac{pe^y}{pe^y + (1-p)e^{-y}} \end{aligned}$$

Example

$$\mathbb{P}(S = 1 | Y = y) = \frac{pe^y}{pe^y + (1-p)e^{-y}}$$

Let us consider specific values of p :

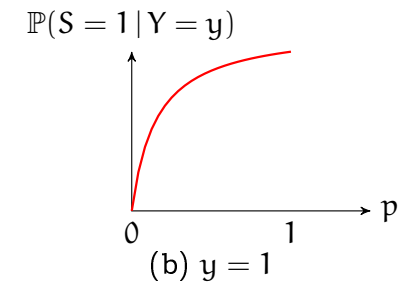
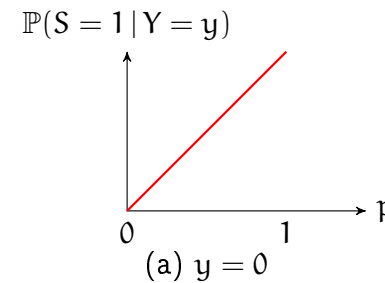


- 1 Expectation, mean, variance
- 2 Common continuous random variables
- 3 Cumulative density function
- 4 Derived distributions
- 5 Joint PDF, conditioning, independence
- 6 Continuous Bayes' rule
- 7 (Advanced Topic) Moment generating function

Example

$$\mathbb{P}(S = 1 | Y = y) = \frac{pe^y}{pe^y + (1-p)e^{-y}}$$

Let us consider specific values of y :



Definition

For a random variable X , define the **moment generating function** $M_X(s)$ (or simply $M(s)$) where

$$M_X(s) = \mathbb{E}(e^{sX}) \quad (50)$$

For continuous random variable, $M(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx$.

Because $\frac{d}{ds} M(s) = \frac{d}{ds} \mathbb{E}(e^{sX}) = \mathbb{E}\left(\frac{d}{ds} e^{sX}\right) = \mathbb{E}(Xe^{sX})$, we have

$$\mathbb{E}(X) = \frac{d}{ds} M(s) \Big|_{s=0} \quad (51)$$

$$\mathbb{E}(X^2) = \frac{d^2}{ds^2} M(s) \Big|_{s=0} \quad (52)$$

(1) Continuous uniform random variable

Continuous uniform random variable over a range $[a, b]$ has PDF

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} M(s) &= \mathbb{E}(e^{sX}) = \int_a^b e^{sx} \frac{1}{b-a} dx \\ &= \frac{1}{s(b-a)} [e^{sx}]_a^b \\ &= \frac{e^{sb} - e^{sa}}{s(b-a)} \end{aligned}$$

(2) Exponential random variable

Exponential random variable with parameter $\lambda > 0$ has PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$M(s) = \int_0^\infty \lambda e^{-\lambda x} e^{sx} dx = \lambda \int_0^\infty e^{(s-\lambda)x} dx = \lambda \frac{e^{(s-\lambda)x}}{s-\lambda} \Big|_0^\infty = \frac{\lambda}{\lambda-s}$$

which requires $s < \lambda$.

$$\mathbb{E}(X) = \frac{d}{ds} \frac{\lambda}{\lambda-s} \Big|_{s=0} = \lambda(-1)(\lambda-s)^{-2}(-1) \Big|_{s=0} = \frac{1}{\lambda}$$

$$\mathbb{E}(X^2) = \frac{d}{ds} \lambda(\lambda-s)^{-2} \Big|_{s=0} = \lambda(-2)(\lambda-s)^{-3}(-1) \Big|_{s=0} = \frac{2}{\lambda^2}$$

$$\text{var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

(1) Continuous uniform random variable

Can use it to compute $\mathbb{E}(X)$ etc.

$$\begin{aligned} \mathbb{E}(X) &= \frac{d}{ds} \left(\frac{e^{sb} - e^{sa}}{s(b-a)} \right) \Big|_{s=0} \\ &= \frac{e^{sb} - e^{sa}}{(b-a)} (-1)s^{-2} + \frac{be^{sb} - ae^{sa}}{s(b-a)} \Big|_{s=0} \\ &= \frac{bse^{sb} - ase^{sa} - e^{sb} + e^{sa}}{s^2(b-a)} \Big|_{s=0} \\ &= \frac{be^{sb} + b^2se^{sb} - a^2se^{sa} - ae^{sa} - be^{sb} + ae^{sa}}{2s(b-a)} \Big|_{s=0} \\ &= \frac{b^2e^{sb} - a^2e^{sa}}{2(b-a)} \Big|_{s=0} = \frac{a+b}{2} \end{aligned}$$

which we used L'Hôpital's rule from 3rd to 4th line.

(3) Normal random variable

Gaussian or **normal** r.v. with parameters ($\mu = 0, \sigma = 1$) has PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\begin{aligned} M(s) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{sx} dx \\ &= \frac{1}{\sqrt{2\pi}} e^{s^2/2} \int_{-\infty}^{\infty} e^{-x^2/2 + sx - s^2/2} dx \\ &= e^{s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-s)^2/2} dx \\ &= e^{s^2/2} \end{aligned}$$

(3) Normal random variable

For $X \sim \mathcal{N}(\mu, \sigma^2)$, let $Y \sim \mathcal{N}(0, 1)$ and therefore $X = \sigma Y + \mu$.

$$M_X(s) = \mathbb{E}\left(e^{s(\sigma Y + \mu)}\right) = e^{s\mu} M_Y(s\sigma) \quad (53)$$

For Gaussian r.v.,

$$M_X(s) = e^{s\mu} e^{(\sigma s)^2/2} = e^{(\sigma^2 s^2/2) + \mu s}$$

(3) Normal random variable

$$\begin{aligned} \mathbb{E}(X) &= \frac{d}{ds} e^{(\sigma^2 s^2/2) + \mu s} \Big|_{s=0} \\ &= (\sigma^2 s + \mu) e^{(\sigma^2 s^2/2) + \mu s} \Big|_{s=0} \\ &= \mu \end{aligned}$$

$$\begin{aligned} \mathbb{E}(X^2) &= \frac{d}{ds} (\sigma^2 s + \mu) e^{(\sigma^2 s^2/2) + \mu s} \Big|_{s=0} \\ &= (\sigma^2 s + \mu)^2 e^{(\sigma^2 s^2/2) + \mu s} + \sigma^2 e^{(\sigma^2 s^2/2) + \mu s} \Big|_{s=0} \\ &= \mu^2 + \sigma^2 \end{aligned}$$

$$\text{var}(X) = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

(3) Normal random variable

If X and Y are independent r.v., and $Z = X + Y$, then

$$M_Z(s) = M_X(s) M_Y(s) \quad (54)$$

Let $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$. Then,

$$\begin{aligned} M_Z(s) &= M_X(s) M_Y(s) = e^{(\sigma_x^2 s^2/2) + \mu_x s} e^{(\sigma_y^2 s^2/2) + \mu_y s} \\ &= e^{((\sigma_x^2 + \sigma_y^2) s^2/2) + (\mu_x + \mu_y) s} \end{aligned}$$

which implies that

$$Z \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

Conclusions

We developed the basic toolset for continuous random variables, and looked at some common distributions.