Several functional forms of $f(x_i)$ are as follows.

Discrete uniform distribution (離散型一様分布):

$$f(x) = \begin{cases} \dfrac{1}{N}, & x = 1, 2, \cdots, N \\ 0, & \text{otherwise} \end{cases}$$

where $N = 1, 2, \cdots$.

Bernoulli distribution (ベルヌイ分布):

$$f(x) = \begin{cases} p^x (1-p)^{1-x}, & x = 0, 1 \\ 0, & \text{otherwise} \end{cases}$$

where $0 \le p \le 1$.

Binomial distribution (二項分布):

$$f(x) = \begin{cases} \dbinom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \cdots, n \\ 0, & \text{otherwise,} \end{cases}$$

where $0 \le p \le 1$ and $n = 1, 2, \cdots$.

$$(a+b)^n = \sum_{x=0}^{n} {}_nC_x a^x b^{n-x} \quad \longrightarrow \quad \text{Binomial Theorem (二項定理)}$$

$${}_nC_x = \binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad n! = 1 \cdot 2 \cdots n \ \text{(factorial of } n\text{)}$$

$X \sim B(n, p)$

Poisson distribution (ポアソン分布):

$$f(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x = 0, 1, \cdots \\ 0, & \text{otherwise,} \end{cases}$$

where $\lambda > 0$.

< Review > e

Note that the definition of e is given by:

$$e = \lim_{x \to 0} (1+x)^{\frac{1}{x}} = \lim_{h \to \infty} \left(1 + \frac{1}{h}\right)^h$$
$$= 2.71828182845905.$$

Notation

$$\exp(x) = e^x$$

< Review > Taylor series expansion about x_0

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \cdots$$

where the kth derivative of $f(x)$ is $f^{(k)}(x)$.

Taylor series expansion of $f(x) = e^x$ about $x = 0$

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

where $f^{(k)}(x) = e^x$. Set $x = \lambda$ and $k = x$.

$$1 = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}$$

Geometric distribution (幾何分布):

$$f(x) = \begin{cases} p(1-p)^x, & x = 0, 1, \cdots \\ 0, & \text{otherwise,} \end{cases}$$

where $0 < p \le 1$.

Negative binomial distribution (負の二項分布):

$$f(x) = \begin{cases} \dbinom{r+x-1}{x} p^r (1-p)^x, & x = 0, 1, \cdots \\ 0, & \text{otherwise,} \end{cases}$$

where $0 < p \le 1$ and $r > 0$.

Taylor series expansion of $f(x) = (1-x)^{-r}$ about $x = 0$

$$f(x) = 1 + rx + \frac{r(r+1)}{2!}x^2 + \frac{r(r+1)(r+2)}{3!}x^3 + \cdots + \binom{r+k-1}{k}x^k + \cdots$$

where $f^{(k)}(x) = \binom{r+k-1}{k}x^k$.

$$(1-x)^{-r} = \sum_{k=0}^{\infty}\binom{r+k-1}{k}x^k$$

Set $x = 1-p$ and $k = x$

$$p^{-r} = \sum_{x=0}^{\infty}\binom{r+x-1}{x}(1-p)^x, \quad \text{i.e.,} \quad 1 = \sum_{x=0}^{\infty}\binom{r+x-1}{x}p^r(1-p)^x$$

Hypergeometric distribution (超幾何分布):

$$f(x) = \begin{cases} \dfrac{\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}, & x = 0, 1, \cdots, n \\ 0, & \text{otherwise,} \end{cases}$$

where $M = 1, 2, \cdots$, $K = 0, 1, \cdots, M$, and $n = 1, 2, \cdots, M$.

In Example 1.2 (Section 1.1), all the possible values of X are 0, 1, 2 and 3. (note that X denotes the number of heads when a coin is cast three times).

That is, $x_1 = 0$, $x_2 = 1$, $x_3 = 2$ and $x_4 = 3$ are assigned in this case.

The probability that X takes x_1, x_2, x_3 or x_4 is given by:

$$\begin{aligned} P(X=0) &= f(0) = P(\{\omega_8\}) = \frac{1}{8}, \\ P(X=1) &= f(1) = P(\{\omega_4, \omega_6, \omega_7\}) \\ &= P(\{\omega_4\}) + P(\{\omega_6\}) + P(\{\omega_7\}) = \frac{3}{8}, \\ P(X=2) &= f(2) = P(\{\omega_2, \omega_3, \omega_5\}) \\ &= P(\{\omega_2\}) + P(\{\omega_3\}) + P(\{\omega_5\}) = \frac{3}{8}, \\ P(X=3) &= f(3) = P(\{\omega_1\}) = \frac{1}{8}, \end{aligned}$$

which can be written as:

$$P(X = x) = f(x) = \frac{3!}{x!(3-x)!}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x}, \quad x = 0, 1, 2, 3.$$

For $P(X = 1)$ and $P(X = 2)$, note that each sample point is mutually exclusive.

The above probability function is called the **binomial distribution (二項分布)**.

Thus, it is easy to check $f(x) \geq 0$ and $\sum_x f(x) = 1$ in Example 1.2.

Continuous Random Variable (連続型確率変数) and Probability Density Function (確率密度関数): Whereas a discrete random variable assumes at most a countable set of possible values, a continuous random variable X takes any real number within an interval I.

For the interval I, the probability which X is contained in A is defined as:

$$P(X \in I) = \int_I f(x)\,\mathrm{d}x.$$

For example, let I be the interval between a and b for $a < b$. Then, we can rewrite $P(X \in I)$ as follows:

$$P(a < X < b) = \int_a^b f(x)\,\mathrm{d}x,$$

where $f(x)$ is called the **probability density function (確率密度関数)** of X, or simply the **density function (密度関数)** of X.

In order for $f(x)$ to be a probability density function, $f(x)$ has to satisfy the following properties:

$$f(x) \geq 0,$$
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$$

Some functional forms of $f(x)$ are as follows:

Uniform distribution (一様分布):

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise,} \end{cases}$$

where $-\infty < a < b < \infty$.

$X \sim U(a, b)$

Normmal distribution (正規分布):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big), \qquad -\infty < x < \infty$$

where $-\infty < \mu < \infty$ and $\sigma > 0$.

$X \sim N(\mu, \sigma^2)$

$N(0, 1)$ = Standard normal distribution

Exponential distribution (指数分布):

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $\lambda > 0$

Gamma distribution (ガンマ分布):

$$f(x) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $\lambda > 0$ and $r > 0$.

Gamma dist. with $r = 1 \iff$ Exponential dist.

Gamma function: $\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} \mathrm{d}x, \quad a > 0$

$\Gamma(a+1) = a\Gamma(a)$ —> Use integration by parts (部分積分)

$\Gamma(n+1) = n!$ for integer n

$\Gamma(n+\frac{1}{2}) = \dfrac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n}\sqrt{\pi}, \quad \Gamma(\frac{1}{2}) = 2\Gamma(\frac{3}{2}) = \sqrt{\pi}$

Beta distribution (ベータ分布):

$$f(x) = \begin{cases} \dfrac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, & 0 < x < 1 \\ 0, & \text{otherwise,} \end{cases}$$

where $a > 0$ and $b > 0$.

Beta function: $$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$B(a,b) = B(b,a)$$

Cauchy distribution (コーシー分布):

$$f(x) = \frac{1}{\pi\beta(1+(x-\alpha)^2/\beta^2)}, \qquad -\infty < x < \infty$$

where $-\infty < \alpha < \infty$ and $\beta > 0$.

Log-normal distribution (対数正規分布):

$$f(x) = \begin{cases} \dfrac{1}{x\sqrt{2\pi\sigma^2}} \exp\Big(-\dfrac{1}{2\sigma^2}(\ln x - \mu)^2\Big), & 0 < x < \infty \\ 0, & \text{otherwise,} \end{cases}$$

where $-\infty < \mu < \infty$ and $\sigma > 0$.

Double exponentioal distribution (二重指数分布), or Laplace distribution (ラプラス分布):

$$f(x) = \frac{1}{2\beta}\exp\Big(-\frac{|x-\alpha|}{\beta}\Big), \qquad -\infty < x < \infty$$

where $-\infty < \alpha < \infty$ and $\beta > 0$.

Weibull distribution (ワイブル分布):

$$f(x) = \begin{cases} abx^{b-1}\exp(-ax^b), & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $a > 0$ and $b > 0$.

Logistic distribution (ロジスティック分布):

$$F(x) = (1 + e^{-(x-\alpha)/\beta})^{-1}, \qquad -\infty < x < \infty$$

where $-\infty < \alpha < \infty$ and $\beta > 0$.

Pareto distribution (パレート分布):

$$f(x) = \begin{cases} \dfrac{\theta x_0^\theta}{x^{\theta+1}}, & x > x_0 \\ 0, & \text{otherwise,} \end{cases}$$

where $x_0 > 0$ and $\theta > 0$.

Gumbel distribution (ガンベル分布), or Extreme value distribution (極値分布):

$$F(x) = \exp(-e^{-(x-\alpha)/\beta}), \qquad -\infty < x < \infty$$

where $-\infty < \alpha < \infty$ and $\beta > 0$.

t distribution (t 分布):

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}\frac{1}{\sqrt{k\pi}}\Big(1 + \frac{x^2}{k}\Big)^{-(k+1)/2}, \qquad -\infty < x < \infty$$

where $k > 0$.

$X \sim t(k)$ —> t dist. with k degrees of freedom (自由度)

$t(1) \iff$ Cauchy dist.

$t(\infty) \iff N(0,1)$

F distribution (F分布):

$$f(x) = \begin{cases} \dfrac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\Big(\dfrac{m}{n}\Big)^{m/2}\dfrac{x^{m/2-1}}{(1+\frac{m}{n}x)^{(m+n)/2}}, & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $m, n = 1, 2, \cdots$.

$X \sim F(m, n)$ —> F dist. with (m, n) degrees of freedom

Chi-square distribution (カイ二乗分布):

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\frac{k}{2})2^{k/2}} x^{k/2-1} e^{-x/2}, & x > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where $k = 1, 2, \cdots$.

$X \sim \chi^2(k)$ —> χ^2 dist. with k degrees of freedom

For a continuous random variable, note as follows:

$$P(X = x) = \int_x^x f(t)\,\mathrm{d}t = 0.$$

In the case of discrete random variables, $P(X = x_i)$ represents the probability which X takes x_i, i.e., $p_i = f(x_i)$.

Thus, the probability function $f(x_i)$ itself implies probability.

However, in the case of continuous random variables, $P(a < X < b)$ indicates the probability which X lies on the interval (a, b).