Random Variable: A random variable is a real-valued function defined on a sample space.

Discrete Random Variable: When the possible values of a random variable are countable, then we say the random variable is a discrete random variable.

It is relatively easier to understand discrete random variables. Note that countable means that either the set has finite many elements or we can establish a one-one correspondence between the elements and the integers $1, 2, 3, \ldots$. So we can list all the elements in the countable set one by one.

For example, an infinite-countable-many set usually has the form

$$\{a_1, a_2, a_3, \ldots, a_k, \ldots, \}.$$

Notation

1. We use ω to denote the outcome in the sample space S.

2. Uppercase letters are usually used to denote random variables. For example, X is a random variable. We use lowercase letter x to denote the value of X.

3. We use $\mathcal{X}$ to denote the collection of the possible values of the random variable X. So

$$\mathcal{X} = \{X(\omega),\ \omega \in S\}.$$

4. If there are more than two random variables, like X and Y, then we add a subscript like $\mathcal{X}_X$ $\mathcal{X}_Y$ to distinguish them.

5. Often we use the abbreviation r.v. instead of random variable.

Discrete Random Variable Defined on a Finite Sample Space

Let us consider a finite sample space.

$$S = \{\omega_1, \omega_2, \ldots, \omega_n\}.$$

Let X be a random variable on S, which means that X is a real-valued function whose arguments are ωs in S.

Let $x_i = X(\omega_i)$ for $i = 1, 2, \ldots, n$. Then

$$\mathcal{X} = \{x_1, x_2, \ldots, x_n\}.$$

Note that it is ok to have some x_is that are the same.

Any well-defined rule can be used to define a random variable X on S. We see some examples in the next few pages.

Examples of Discrete R.V.

Ex1: Toss a coin. $S = \{H, T\}$.

Let $X(T) = 1$ and $X(T) = 0$ then we define a r.v. X. $\mathcal{X} = \{1, 0\}$.

Ex2: Roll a die. $S = \{1, 2, 3, 4, 5, 6\}$.

Let $X(i) = i$ for $i = 1, 2, 3, 4, 5, , 6$. $\mathcal{X}_X = \{1, 2, 3, 4, 5, 6\}$.

Or we can define another r.v. Y by $Y(1) = Y(3) = Y(5) = 1$ and $Y(2) = Y(4) = Y(6) = 0$. $\mathcal{X}_Y = \{1, 0\}$. In this example we see different arguments could give the same value to Y.

Ex3: Toss 2 coins.

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

Let X be the number of heads. Then $X((H, H)) = 2$, $X((H, T)) = X((T, H)) = 1$, $X((T, T)) = 0$. $\mathcal{X} = \{2, 1, 0\}$.

Probability Function

Since the argument of the random variable X is random, the value of X is random too. It is natural to ask what is the probability that the value of X is equal to x? (x is some real number).

Probability Function: Probability function (p.f.) of the discrete random variable X is defined as

$$f(x) = \text{Pr}(X = x) = \sum_{\omega : X(\omega) = x} p(\omega).$$

$p(\omega)$ is the probability distribution on the finite sample space.

We use probability function $f(x)$ to describe the distribution of X.

If A is a subset of $\mathcal{X}$ then

$$\text{Pr}(X \in A) = \sum_{x_i \in A} f(x_i).$$

In Ex1, $f(1) = \Pr(X = 1) = p(H) = 0.5$ and $f(0) = \Pr(X = 0) = p(T) = 0.5$.

In Ex2, $\Pr(X = i) = p(i) = \frac{1}{6}$. $i = 1, 2, 3, 4, 5, 6$.

$$\Pr(Y = 1) = p(1) + p(3) + p(5) = \frac{1}{2},$$

$$\Pr(Y = 0) = p(2) + p(4) + p(6) = \frac{1}{2}.$$

In Ex3, $f(2) = \Pr(X = 2) = p((H, H)) = \frac{1}{4}$ and $f(1) = \Pr(X = 1) = p((H, T)) + p((T, H)) = \frac{1}{4} + \frac{1}{4} = 1/2$, $f(0) = \Pr(X = 0) = p((T, T)) = \frac{1}{4}$.

Ex4: Roll 2 dice. Let X be the sum of the two numbers. Find $\Pr(X = 5)$? Solution: $f(5) = \Pr(X = 5) = p((4, 1)) + p((1, 4)) + p((3, 2)) + p((2, 3)) = \frac{4}{36} = \frac{1}{9}$.

Three Common Discrete Distributions

Let $x_1, x_2, \ldots, x_k$ be distinct real numbers. If we have a function $f(x)$ such that

$$f(x_i) \geq 0,$$

for all $i = 1, 2, \ldots, n$, and

$$\sum_{i=1}^{n} f(x_i) = 1.$$

Then f can be a probability function of X which takes value from $\mathcal{X} = \{x_1, x_2, \ldots, x_k\}$.

Ex1: X takes value from $\{1, 0\}$. Let $\Pr(X = 1) = f(1) = p$ and $\Pr(X = 0) = f(0) = 1 - p = q$. Then We say X is a Bernoulli random variable whose probability of success is p. We write $X \sim Ber(p)$.

Ex2: Let $\mathcal{X} = \{0, 1, 2, \ldots, n\}$. Take a positive real number p, $0 < p < 1$. Let

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

By Binomial theorem we check

$$\sum_{k=0}^{n} f(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$

Thus we can use $\{f(k)\}$ to define a distribution. We say X is a binomial random variable with parameters (n, p) if X takes value from $\{0, 1, 2, \ldots, n\}$ and $\Pr(X = k) = f(k) = \binom{n}{k} p^k (1-p)^{n-k}$. We write $X \sim Bin(n, p)$.

Ex3: We say X is a uniform random variable on $\{x_1, x_2, x_3, \ldots, x_k\}$ if $\Pr(X = x_i) = f(x_i) = \frac{1}{k}$ for $i = 1, 2, \ldots, k$.

Remark: Two random variables can have the same distribution.

Some Bernoulli examples: 1. In a state governor election voters vote for two candidates, R and D. Randomly pick a voter, let $X = 1$ if the voter votes for R and $X = 0$ if the voter votes for D. Then X is a Bernoulli random variable. What is the probability of success p? p should be the probability that a randomly selected voter votes for R.

$$p = \frac{\#\ of\ voters\ for\ R}{\#\ of\ all\ voters}.$$

2. Let A be some event. We define a random variable X as follows

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

X is a Bernoulli random variable. Its probability of success is

$$\begin{aligned} p &= \Pr(X = 1) \\ &= \sum_{\omega_i : X(\omega_i = 1)} p(\omega_i) \\ &= \sum_{\omega \in A} p(\omega_i) = p(A). \end{aligned}$$

Some Binomial examples: 1. Toss a biased coin 100 times $p(H) = 0.6$ $p(T) = 0.4$. Let X be the number of heads in 100 tosses.

$$\Pr(X = k) = \binom{100}{k} 0.6^k 0.4^{100-k}.$$

So $X \sim Bin(100, 0.6)$.

2. Let X be the number of women in a Jury. $n = 12$. X takes value from $\{0, 1, 2, \ldots, 12\}$. Let p be the probability that a random chosen person is a woman and $1 - p$ be the probability that a random chosen person is a man. Then

$$\Pr(X = k) = \binom{12}{k} p^k (1 - p)^{12-k}.$$

So $X \sim Bin(12, p)$.

What is p? p is the percentage of women in the district.

Some Uniform r.v. examples: 1. Roll a die. $f(k) = \frac{1}{6}$, $k = 1, 2, 3, 4, 5, 6$.

2. The last digit of your SSN. It is uniformly distributed on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Normalizing constant: Suppose that X has probability function

$$f(x) = \frac{c}{x^2}, \ x = 1, 2, 3, 4, 5$$

Find the value of the constant c.

Solution: Since f is a p.f. we must have

$$1 = \sum_x f(x) = \sum_{x=1}^{5} \frac{c}{x^2}.$$

Thus $c = (\sum_{x=1}^{5} \frac{1}{x^2})^{-1} = 0.683$.

Remark: c is often called the normalizing constant of the distribution. In some applications we know the probability function upto a constant. Then a practically important question is to compute the value of that constant, as we did in the above example.

Distribution Functions

We start with the discrete random variables.

Definition: Let X be a discrete random variable. $\mathcal{X} = \{x_1, x_2, x_3, \ldots,\}$ and $x_1 < x_2 < x_3 < \ldots$. The distribution function of X is given by

$$F(x) = \text{Pr}(X \leq x) \qquad for \;\; -\infty < x < \infty$$

Note that in terms of the probability function of X

$$F(x) = \sum_{x_i : x_i \leq x} f(x_i).$$

Sometimes we also call $F(x)$ the CDF of X.

Example

Suppose $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$. $f(x_1) = \frac{1}{8}$, $f(x_2) = \frac{2}{8}$, $f(x_3) = \frac{2}{8}$, $f(x_4) = \frac{3}{8}$. f is the p.f. of X. Let's compute the distribution function of X.

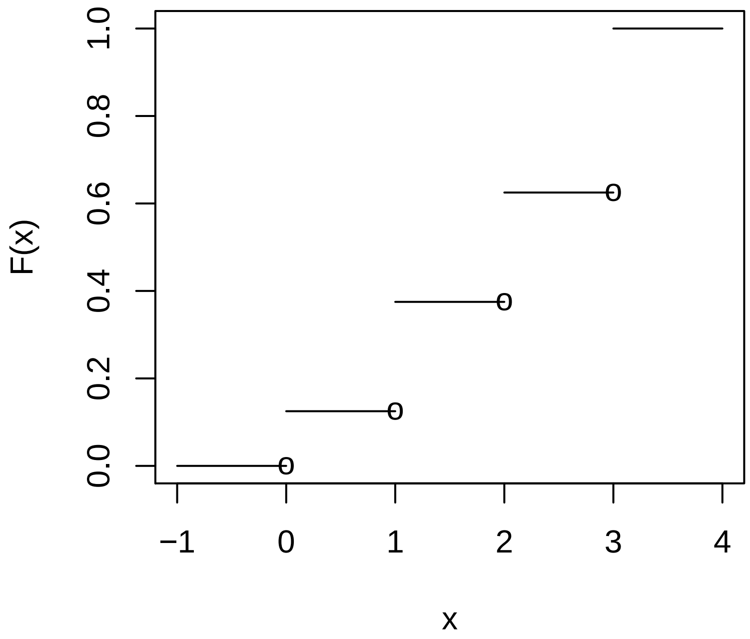
$F(-1) = 0$.

$F(0) = f(x_1) = \frac{1}{8}$, $F(0.5) = f(x_1) = \frac{1}{8}$.

$F(1) = f(x_1) + f(x_2) = \frac{3}{8}$, $F(1.5) = f(x_1) + f(x_2) = \frac{3}{8}$.

$F(2) = f(x_1) + f(x_2) + f(x_3) = \frac{5}{8}$, $F(2.5) = f(x_1) + f(x_2) + f(x_3) = \frac{5}{8}$.

$F(3) = f(x_1) + f(x_2) + f(x_3) + f(x_4) = 1$, $F(3.25) = f(x_1) + f(x_2) + f(x_3) + f(x_4) = 1$.



The above figure shows the sketch $F(X)$. Form the graph we can see that $F(x)$ has three properties. If $x_1 < x_2$ then $F(x_1) \leq F(x_2)$. "=" can be taken. In the graph $F(1) = F(1.5) = 3/8$.

1. $F(x)$ is nondecreasing as x increases.

2. $F(x)$ is continuous from the right at every x. $F(x) = \lim_{\delta>0,\delta\to 0} F(x+\delta)$.

3. $\lim_{x\to -\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$.

Continuous Random Variable and pdf

Definition: X is a continuous random variable if there exists a nonnegative function $f(x)$ defined for all $x \in (-\infty, \infty)$ such that for every set A in the real line $\mathcal{R}$

$$\Pr(X \in A) = \int_A f(x)dx$$

$f(x)$ is called the probability density function (pdf) of X.

On the other hand, if we have a function $f(x)$ such that $f(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} f(x)dx = 1$, then $f(x)$ is a legitimate pdf of a continuous random variable.

Remarks:

1. connection between CDF and pdf:

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad f(x) = \frac{dF(x)}{dx}.$$

2. $P(a \leq X \leq b) = \int_a^b f(x)dx$, $P(X = a) = \int_a^a f(x)dx = 0$.

Ex1: Uniform distribution on (a, b).

We say X is a uniform random variable on the interval (a, b) if its pdf is given by

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

We write $X \sim Unif(a, b)$.

We check (1). $f(x) \geq 0$ and (2).

$$\begin{aligned} & \int_{-\infty}^{\infty} f(x)dx \\ = & \int_{-\infty}^{a} 0dx + \int_{a}^{b} \frac{1}{b-a}dx + \int_{b}^{\infty} 0dx \\ = & 0 + 1 + 0 = 1. \end{aligned}$$

Therefore we know the defined $f(x)$ is a legitimate pdf.

Ex2: Suppose $X \sim Unif(0, 10)$. Find (a). $P(X < 3)$, (b). $P(X > 7)$, and (c). $P(1 < X < 6)$.

Solution: Since $X \sim Unif(0, 10)$, we know

$$f(x) = \begin{cases} \frac{1}{10} & 0 < x < 10 \\ 0 & \text{otherwise} \end{cases}$$

(a).

$$P(X < 3) = \int_{-\infty}^{3} f(x)dx = \int_{0}^{3} \frac{1}{10}dx = 0.3$$

(b).

$$P(X > 7) = \int_{7}^{\infty} f(x)dx = \int_{7}^{10} \frac{1}{10}dx = 0.3$$

(c).

$$P(1 < X < 6) = \int_{1}^{6} f(x)dx = \int_{1}^{6} \frac{1}{10}dx = 0.5$$

Ex3: The pdf of a r.v. X is

$$f(x) = \begin{cases} cx & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of c and compute $P(1 \le X \le 2)$.

Solution: $f(x)$ must satisfy the following condition

$$1 = \int_{-\infty}^{\infty} f(x)dx$$

thus we establish the following equation

$$1 = \int_0^4 cxdx = \frac{1}{2}cx^2|_0^4 = \frac{1}{2}c4^2 - \frac{1}{2}c0^2 = 8c$$

So $c = \frac{1}{8}$. Then we have

$$\begin{aligned} P(1 \le X \le 2) &= \int_1^2 \frac{x}{8}dx \\ &= \frac{1}{16}x^2|_1^2 \\ &= \frac{1}{16}2^2 - \frac{1}{16}1^2 \\ &= \frac{3}{16}. \end{aligned}$$

Ex4: If a r.v. X has CDF

$$F(x) = \begin{cases} 1 - e^{-x} & 0 \le x \\ 0 & \text{otherwise} \end{cases}$$

Find its pdf and $P(1 \le X \le 2)$.

Solution: Using $f(x) = \frac{dF(x)}{dx}$ we compute

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

$$\begin{aligned} P(1 \le X \le 2) &= \int_1^2 e^{-x} dx \\ &= (-e^{-x})|_1^2 \\ &= -e^{-2} - (-e^{-1}) \\ &= e^{-1} - e^{-2}. \end{aligned}$$

Or we can directly use CDF

$$\begin{gathered} P(1 \le X \le 2) = P(X \le 2) - P(X \le 1) \\ = F(2) - F(1) \\ = (1 - e^{-2}) - (1 - e^{-1}) \\ = e^{-1} - e^{-2}. \end{gathered}$$

Ex5: If $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$ for $-\infty < x < \infty$. Find $F(x)$ and $P(-1 < X < 1)$. Note that $f(x)$ is the pdf of a Cauchy distribution.

Solution:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \int_{-\infty}^{x} \frac{1}{\pi}\frac{1}{1+t^2}dt \\ &= \frac{1}{\pi}\text{arctan}(t)|_{-\infty}^{x} \\ &= \frac{1}{\pi}\text{arctan}(x) - \frac{1}{\pi}\text{arctan}(-\infty) \\ &= \frac{1}{\pi}\text{arctan}(x) + \frac{1}{2}. \end{aligned}$$

$$\begin{aligned} P(-1 < X < 1) &= F(1) - F(-1) \\ &= \frac{1}{\pi}\text{arctan}(1) - \frac{1}{\pi}\text{arctan}(-1) \\ &= \frac{1}{\pi}\frac{\pi}{4} - \frac{1}{\pi}\frac{-\pi}{4} \\ &= \frac{1}{2}. \end{aligned}$$

Ex6: Standard normal distribution. We say X is a standard normal random variable if X has a pdf as follows

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

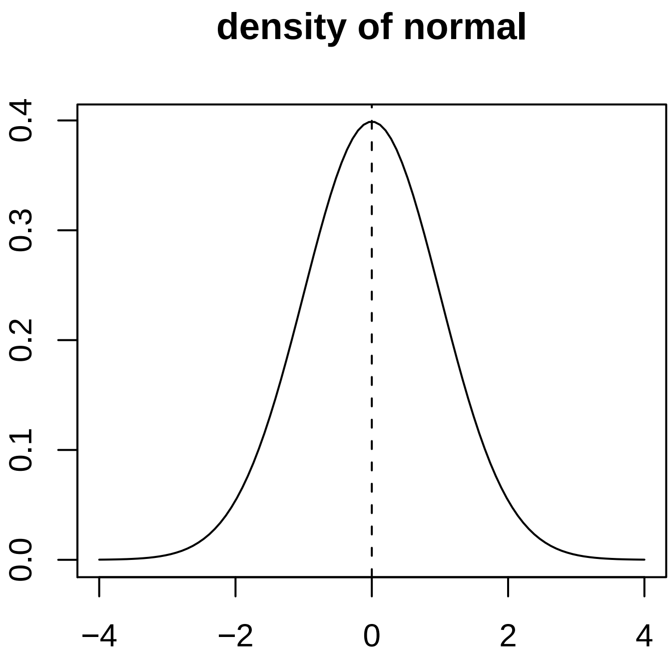
We write $X \sim N(0,1)$. Let's check $f(x)$ is a pdf function. Obviously $f(x) > 0$ for all x. We only need to verify $\int_{-\infty}^{\infty} f(x)dx = 1$. Let $I = \int_{-\infty}^{\infty} f(x)dx$.

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} f(x)dx \int_{-\infty}^{\infty} f(y)dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy \end{aligned}$$

We now change the variables from (x, y) to (r, θ) be letting $y = r\sin(\theta)$, $x = r\cos(\theta)$ then $x^2 + y^2 = r^2$ and

$$\begin{aligned} I^2 &= \int_0^{\infty}\int_0^{2\pi} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta \\ &= \int_0^{2\pi} d\theta \frac{1}{2\pi}(-e^{-\frac{r^2}{2}})|_0^{\infty} = \int_0^{2\pi} d\theta \frac{1}{2\pi} = 1. \end{aligned}$$

So $I = 1$.



The above figure shows the plot of the density function of standard normal.

Normal distribution has countless applications. It is perhaps the most important and useful distribution. We will study its properties in the later part of this course.

Compute the Density of a Transformed Random Variable

Suppose X is a r.v. whose pdf is $f(x)$ and we know $f(x)$. We can construct a new r.v. Y from X through a function g, that is, $Y = g(X)$. For example, $X \sim Unif(0,1)$ and $Y = g(X) = \sqrt{X}$. Then the question is how do we compute the pdf of the new r.v. Y?

To handle this question we consider two kinds of transformations.

Case 1. There is an inverse function g^{-1} such that $X = g^{-1}(Y)$.

Case 2. The inverse function of g does not exist.

Case 1: inverse function exists

Theorem: Let $Y = g(X)$ and $f_X(x)$ is the pdf of X. Suppose g^{-1} exists and g^{-1} is differentiable. Then the pdf of Y is $f_Y(y)$, given by

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|.$$

Ex1: $X \sim Unif(0,1)$ and $Y = g(X) = \sqrt{X}$.

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

In this example, the inverse function g^{-1} exists and $X = Y^2$, $g^{-1}(y) = y^2$. $\frac{dg^{-1}(y)}{dy} = 2y$. Using the theorem,

$$f_Y(y) = f_X(g^{-1}(y))2y$$

$$f_X(g^{-1}(y)) = \begin{cases} 1 & 0 < g^{-1}(y) < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & 0 < y^2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Ex2: $X \sim Unif(0,1)$ and $Y = g(X) = -\frac{\log(X)}{2}$.

In this case, g^{-1} exists and $X = e^{-2Y}$, $g^{-1}(y) = e^{-2y}$. $\frac{dg^{-1}(y)}{dy} = -2e^{-2y}$. Using the theorem,

$$f_Y(y) = f_X(g^{-1}(y))|-2e^{-2y}|$$

$$f_X(g^{-1}(y)) = \begin{cases} 1 & 0 < g^{-1}(y) < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 2e^{-2y} & 0 < e^{-2y} < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 2e^{-2y} & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Case 2: inverse function does not exist

Ex3: $X \sim N(0,1)$ and $Y = X^2$. In this case we use CDF to compute the pdf.

Y takes value from $[0, \infty)$. Pick any $y \geq 0$,

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

$$\begin{aligned} \frac{dF_Y(y)}{dy} &= \frac{d(F_X(\sqrt{y}) - F_X(-\sqrt{y}))}{dy} \\ &= f_X(\sqrt{y})\frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\frac{-1}{2\sqrt{y}} \\ &= \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] \end{aligned}$$

Note that we used the chain rule to compute the derivative. Substituting $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ into the above equation,

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}} \quad y \geq 0.$$

Discrete Joint Distributions

Joint Probability Function: Consider a pair of discrete random variables (X, Y). For every point (x, y) we let

$$f(x, y) = \Pr(X = x, Y = y).$$

We say $f(x, y)$ is the joint p.f. of (X, Y).

Clearly, if x_0 or y_0 is not one of the possible values of X or Y then $f(x_0, y_0) = 0$.

Let $\mathcal{X}$ and $\mathcal{Y}$ denote the all possible values of X and Y, respectively. Then we must have the identity:

$$1 = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} f(x, y).$$

For each subset A in the xy-plane,

$$\Pr((X, Y) \in A) = \sum_{(x,y) \in A} f(x, y)$$

Example: a table of probabilities

Suppose the joint p.f. of (X, Y) is specified in the following table

	Y				
X	0	1	2	3	4
0	0.08	0.07	0.06	0.01	0.01
1	0.06	0.10	0.12	0.05	0.02
2	0.05	0.06	0.09	0.04	0.03
3	0.02	0.03	0.03	0.03	0.04

Determine (a) $P(X = 2)$; (b) $P(X = Y)$; (c) $P(X \leq 2, Y \leq 2)$; (d) $P(Y = 3)$.

Solution:

$$\begin{aligned} P(X = 2) &= \sum_{x=2,y} f(x, y) = \sum_{y=0}^{4} f(2, y) \\ &= 0.05 + 0.06 + 0.09 + 0.04 + 0.03 \\ &= 0.27. \end{aligned}$$

$$
\begin{aligned}
P(X = Y) &= \sum_{x=y} f(x,y) = \sum_{i=0}^{3} f(i,i) \\
&= 0.08 + 0.10 + 0.09 + 0.03 \\
&= 0.30.
\end{aligned}
$$

$$
\begin{aligned}
P(X \le 2, Y \le 2) &= \sum_{x \le, y \le 2} f(x,y) \\
&= \sum_{x=0}^{2} \sum_{y=0}^{2} f(x,y) \\
&= 0.08 + 0.07 + 0.06 \\
&\quad +0.06 + 0.10 + 0.12 \\
&\quad +0.05 + 0.06 + 0.09 \\
&= 0.69.
\end{aligned}
$$

$$
\begin{aligned}
P(Y = 3) &= \sum_{x, y=3} f(x,y) = \sum_{x=0}^{3} f(x,3) \\
&= 0.01 + 0.05 + 0.04 + 0.03 \\
&= 0.13.
\end{aligned}
$$

Continuous Joint Distributions

Joint Probability Density Function: Consider a pair of continuous random variables (X, Y). The joint pdf of (X, Y) is a bivariate function $f(x, y)$ satisfying two conditions:

1. $f(x, y) \geq 0$

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Let A be some region, and denote $P(A) = P((X, Y) \in A)$.

$$P(A) = \iint_A f(x, y) dx dy.$$

Ex.1

The joint pdf of (X, Y) is given by

$$f(x,y) = \begin{cases} 1 & \text{if } 0 < x < 1, 0 < y < 1 \\ 0 & \text{Otherwise} \end{cases}$$

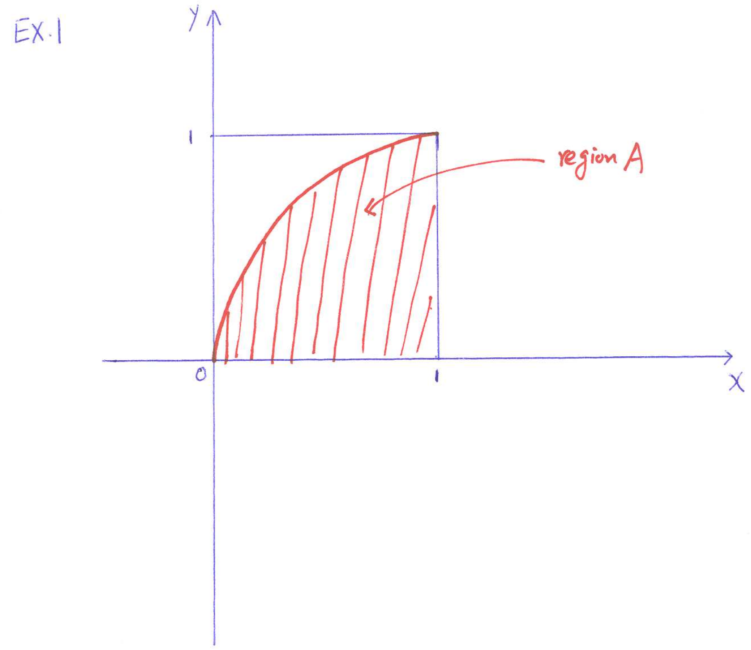
Compute $\Pr((X-1)^2 + Y^2 \leq 1)$.

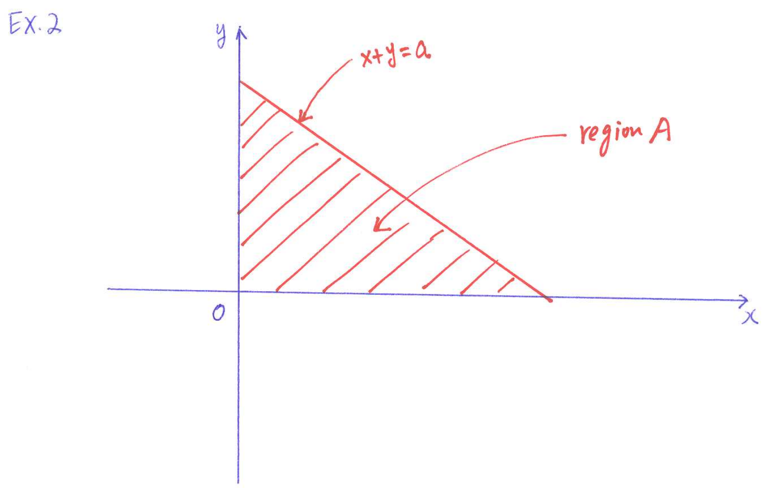
Solution: Let $A = \{(x,y) : (x-1)^2 + y^2 \leq 1\}$

$$\begin{aligned} \Pr((X-1)^2 + Y^2 \leq 1) &= \iint_A f(x,y)dxdy \\ &= \iint_{A \cap [0,1]^2} 1dxdy \\ &= \text{Area of}(\text{A} \cap [0,1]^2) \\ &= \frac{\pi}{4} \end{aligned}$$

$A \cap [0,1]^2$ means the intersection of A and $[0,1]^2$. Similarly, $P(0 < X < 0.5) = 0.5$. (check this yourself)

Geometry and symmetry are very useful when dealing with uniform distributions.





The above figure shows the region A for the double integral in Example 1 and Example 2.

Ex.2

The joint pdf of (X, Y) is given by

$$f(x,y) = \begin{cases} e^{-x-y} & \text{if } 0 < x, 0 < y \\ 0 & \text{Otherwise} \end{cases}$$

Find $P(X + Y \leq a)$.

Solution:

$$\begin{aligned} P(X+Y \leq a) &= \iint_{x+y \leq a} f(x,y)dxdy \\ &= \int_0^a dx[\int_0^{a-x} dye^{-x-y}] \\ &= \int_0^a dx[e^{-x} - e^{-a}] \\ &= (-e^{-x} - e^{-a}x)|_0^a \\ &= 1 - e^{-a}(1+a) \end{aligned}$$

We often need to compute the double integral by a process of repeated single integrations.

Question: Let $Z = X + Y$, what is the CDF of Z? and what is its pdf? Hint: $P(Z \leq a) = P(X + Y \leq a)$.

Ex.3

Suppose the joint pdf of (X, Y) is given by

$$f(x, y) = \begin{cases} 4xy & \text{if } 0 < x < 1, 0 < y < 1 \\ 0 & \text{Otherwise} \end{cases}$$

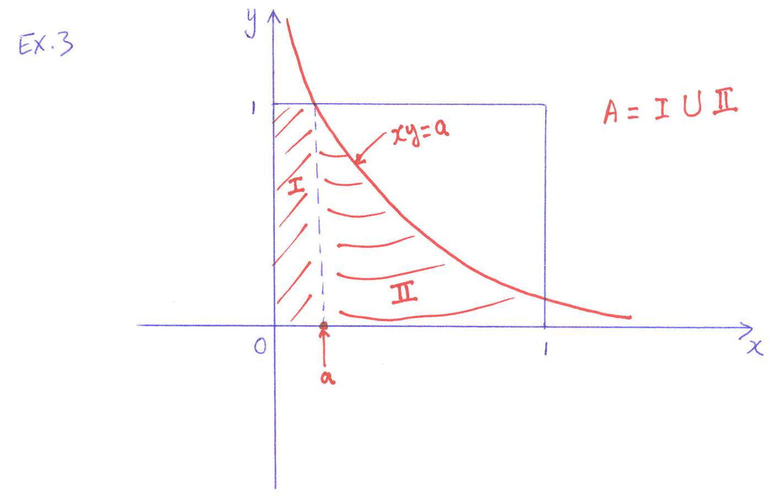
Let $Z = XY$. Find $P(Z \le a)$ and pdf of Z.

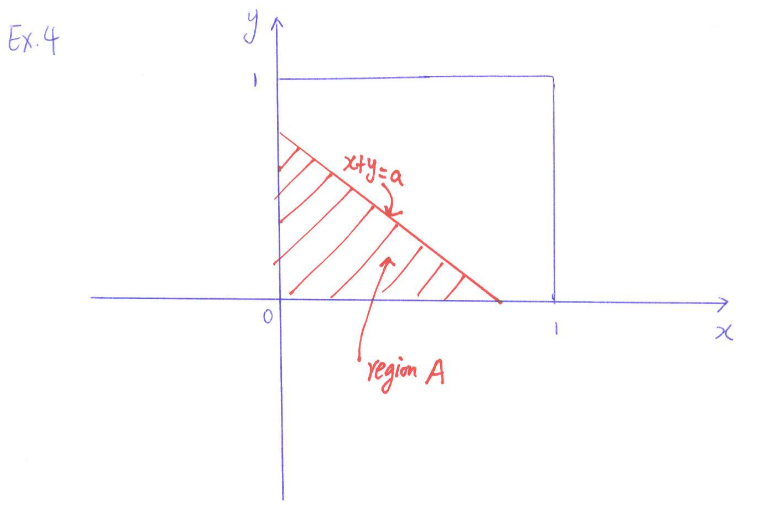
Solution:

$$\begin{aligned} P(Z \le a) &= \iint_{0<x<1,0<y<1,xy\le a} 4xydxdy \\ &= \int_0^a dx[\int_0^1 dy4xy] + \int_a^1 dx[\int_0^{\frac{a}{x}} dy4xy] \\ &= \int_0^a dx[2x] + \int_a^1 dx[\frac{2a^2}{x}] \\ &= a^2 - 2a^2 \ln(a) \end{aligned}$$

Since $F_Z(a) = P(Z \le a)$

$$\begin{aligned} f_Z(a) &= \frac{dP(Z \le a)}{da} \\ &= -4a \ln(a) \end{aligned}$$





The above figure shows the region A for the double integral in Example 3 and Example 4.

Ex.4

Suppose (X, Y) has the joint pdf

$$f(x,y) = \begin{cases} x+y & \text{if } 0<x<1, 0<y<1 \\ 0 & \text{Otherwise} \end{cases}$$

Let $Z = X + Y$. Find $F_Z(a) = P(Z \le a)$ for $0 < a < 1$.

Solution:

$$\begin{aligned} P(X+Y \le a) &= \iint_{0<x,0<y,x+y\le a} x + y dx dy \\ &= \int_0^a dx[\int_0^{a-x} dy \ (x+y)] \\ &= \int_0^a dx[(a-x)x + \frac{1}{2}(a-x)^2] \\ &= \int_0^a dx[\frac{a^2 - x^2}{2}] \\ &= [\frac{a^2}{2}x - \frac{1}{6}x^3]|_0^a = \frac{a^3}{3}. \end{aligned}$$

Bivariate Distribution Functions

Definition: The joint distribution function or joint d.f. (or CDF) of (X, Y) is defined as

$$F(x, y) = \text{Pr}(X \le x, Y \le y),$$

$-\infty < x < \infty,\ -\infty < y < \infty.$

Properties: If (X, Y) has a continuous joint p.d.f $f(x, y)$ then

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(r, s) dr ds.$$

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

$$\begin{aligned} F_X(x) &= \text{Pr}(X \le x) \\ &= \lim_{y \to \infty} \text{Pr}(X \le x, Y \le y) \\ &= \lim_{y \to \infty} F(x, y). \end{aligned}$$

$$\textit{Likewise,}\ F_Y(y) = \lim_{x \to \infty} F(x, y).$$

Ex.5

Suppose that the joint p.d.f of (X, Y) is

$$f(x,y) = \begin{cases} \frac{1}{8}(x+y) & \text{if } 0 < x < 2, 0 < y < 2 \\ 0 & \text{Otherwise} \end{cases}$$

Compute (a). $F(x, y)$ and (b). $\Pr(0.5 < X < 1.5, 0.5 < Y < 1.5)$.

Solution: (a). We first consider the case when $0 < x < 2$ and $0 < y < 2$.

$$\begin{aligned} F(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(r,s)drds \\ &= \int_0^y \int_0^x \frac{1}{8}(r+s)drds \\ &= \int_0^y [\frac{1}{8}(\frac{1}{2}r^2 + sr)|_0^x]ds \\ &= \int_0^y [\frac{1}{16}x^2 + \frac{1}{8}sx]ds \\ &= \frac{1}{16}x^2y + \frac{1}{16}y^2x \\ &= \frac{1}{16}xy(x+y). \end{aligned}$$

If $x \geq 2, y \geq 2,\ F(x,y) = P(X \leq x, Y \leq y) = 1$.

If $x < 0, y < 0,\ F(x,y) = P(X \leq x, Y \leq y) = 0$.

If $0 < x < 2, y \geq 2$,

$$\begin{aligned} F(x,y) &= P(X \leq x, Y \leq y) \\ &= P(X \leq x, Y \leq 2) \\ &= F(X, 2) \\ &= \frac{1}{16}x2(x+2) = \frac{1}{8}x(x+2). \end{aligned}$$

If $x \geq 2, 0 < y < 2$,

$$\begin{aligned} F(x,y) &= P(X \leq x, Y \leq y) \\ &= P(X \leq 2, Y \leq y) \\ &= F(2, y) \\ &= \frac{1}{16}2y(2+y) = \frac{1}{8}y(y+2). \end{aligned}$$

Solution (b).

$$\begin{aligned}
& \text{Pr}(0.5 < X < 1.5, 0.5 < Y < 1.5) \\
= \; & P(0.5 < X < 1.5, Y < 1.5) \\
& -P(0.5 < X < 1.5, Y < 0.5) \\
= \; & [P(X < 1.5, Y < 1.5) - P(X < 0.5, Y < 1.5)] \\
& -[P(X < 1.5, Y < 0.5) - P(X < 0.5, Y < 0.5)] \\
= \; & [F(1.5, 1.5) - F(0.5, 1.5)] - \\
& [F(1.5, 0.5) - F(0.5, 0.5)]
\end{aligned}$$

By (a) we can compute

$$F(1.5, 1.5) = \frac{1}{16} 1.5 \times 1.5 (1.5 + 1.5) = 0.421875$$

$$F(1.5, 0.5) = \frac{1}{16} 1.5 \times 0.5 (1.5 + 0.5) = 0.09375$$

$$F(0.5, 1.5) = \frac{1}{16} 0.5 \times 1.5 (0.5 + 1.5) = 0.09375$$

$$F(0.5, 0.5) = \frac{1}{16} 0.5 \times 0.5 (0.5 + 0.5) = 0.015625$$

$$\text{Pr}(0.5 < X < 1.5, 0.5 < Y < 1.5) = 0.25.$$