ECE541: Stochastic Signals and Systems Fall 2019

Lecture 2 - September 12, 2019

Dr. Salim El Rouayheb *Scribe: Peiwen Tian, Lu Liu, Ghadir Ayache*

Chapter 2: Random Variables

Example 1. *Tossing a fair coin twice:*

$$\Omega = \{HH, HT, TH, TT\}.$$

Define for any $\omega \in \Omega$, $X(\omega)$=number of heads in ω. $X(\omega)$ is a random variable.

Definition 1. *A random variable (RV) is a function X: $\Omega \to \mathbb{R}$.*

Example 2. *Let w be the temperature in °F at 3:00 pm on Thursday afternoon. Let X be the r.v. which the temperature in °C. Then*

$$X = \frac{5}{9}(w - 32)$$

Definition 2 (Cumulative distribution function(CDF))**.**

$$F(x) = P(X \leq x). \tag{1}$$

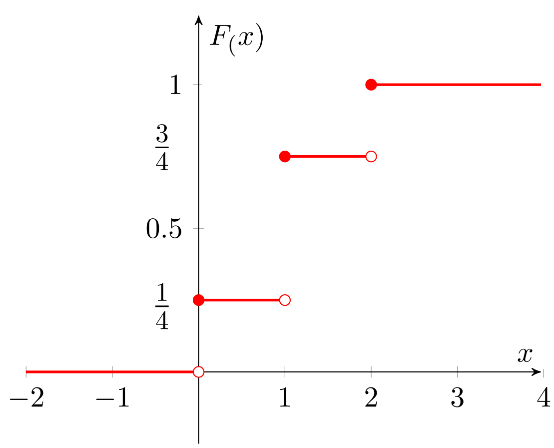

Figure 1: Cumulative distribution function of x

Example 3. *The cumulative distribution function of x is as (Figure 1)*

$$F_X(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{4} & 0 \leq x < 1, \\ \frac{3}{4} & 1 \leq x < 2, \\ 1 & x \geq 2. \end{cases}$$

Lemma 1. *Properties of CDF*

(1)

$$\lim_{x\to-\infty} F_X(x) = 0 \quad (2)$$
$$\lim_{x\to+\infty} F_X(x) = 1, \quad (3)$$

(2) $F_X(x)$ *is non-decreasing:*

$$x_1 \le x_2 \implies F_X(x_1) \le F_X(x_2) \quad (4)$$

(3) $F_X(x)$*is continuous from the right*

$$\lim_{\epsilon\to 0} F_X(x+\epsilon) = F_X(x), \epsilon > 0 \quad (5)$$

(4)

$$P(a \le X \le b) = P(X \le b) - P(X \le a) + P(X = a) \quad (6)$$
$$= F_X(b) - F_X(a) + P(X = a) \quad (7)$$

(5)

$$P(X = a) = \lim_{\varepsilon\to 0} F_X(a) - F_X(a-\varepsilon), \varepsilon > 0 \quad (8)$$

Definition 3. *If random variable X has finite or countable number of values, X is called discrete. If It is uncountable, then it is continuous.*

Remark 1. *A set S is countable if the its elements can be indexed, i.e., we can find a injective function from S to the natural numbers*

Example 4. *Non-countable example:*$\mathbb{R}$.

Example 5. *Countable example: The number of tosses we need till get a Head*

Lemma 2. *If X is continuous, then $F_X(x)$ is continuous.*

Definition 4 (Probability density function(pdf))**.**

$$f_X(x) = \frac{dF_X(x)}{dx} \; (X \text{ is continuous}). \quad (9)$$

Example 6. *Gaussian random variable: Normal/ Gaussian Distribution.*

By definition,

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

Therefore,

$$F_X(a) = P(x \le a) = \int_{-\infty}^{a} f_X(x)dx,$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx.$$

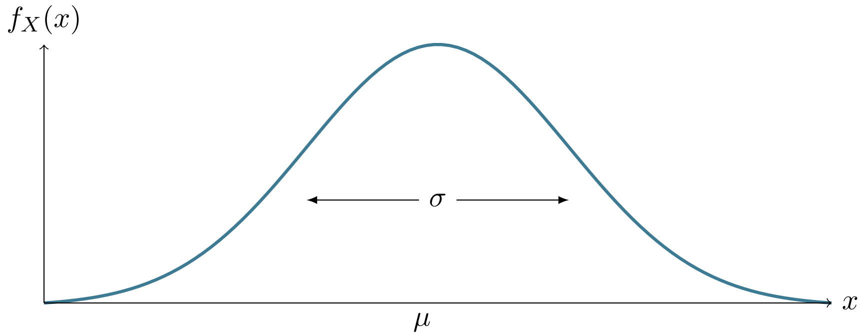

Figure 2: Gaussian distribution pdf.

We should always have:

$$\int_{-\infty}^{+\infty} f_X(x)dx = 1.$$

Definition 5 (mean, variance of a RV X)**.** *For the continuous case:*

$$\begin{aligned} E(X) = \mu &= \int_{-\infty}^{+\infty} x f_X(x)dx, \\ V(X) = \sigma^2 &= \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)dx. \end{aligned}$$

For the discrete case:

$$\begin{aligned} E(X) &= \mu = \sum_{i=-\infty}^{+\infty} x_i P(X = x_i), \\ V(X) &= \sigma^2 = \sum_{i=-\infty}^{+\infty} (x_i - \mu)^2 P(X = x_i). \end{aligned}$$

Example 7. *X is uniformly distributed in* $[0, 1]$.

$$F_X(x) = \begin{cases} 0 & x < 0, \\ \int_0^x 1dx = x & 0 \leq x < 1, \\ 1 & x \geq 1. \end{cases}$$

$$\begin{aligned} E(X) &= \int_0^1 X \times 1dx = \frac{1}{2}, \\ V(X) &= \int_0^1 (X - \frac{1}{2})^2 \times 1dx = \frac{1}{12}. \end{aligned}$$

Lemma 3 (Probability Density Functions)**.**

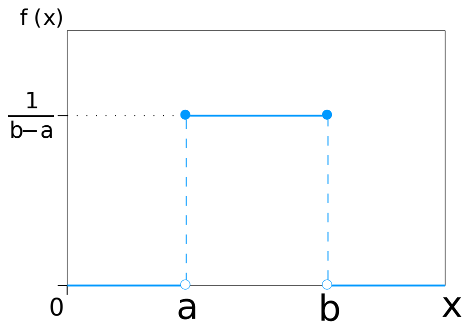

Figure 3: Uniform distribution.[1]

(1) Uniform X uniform over [a,b]:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

(2) Gaussian distribution:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \tag{11}$$

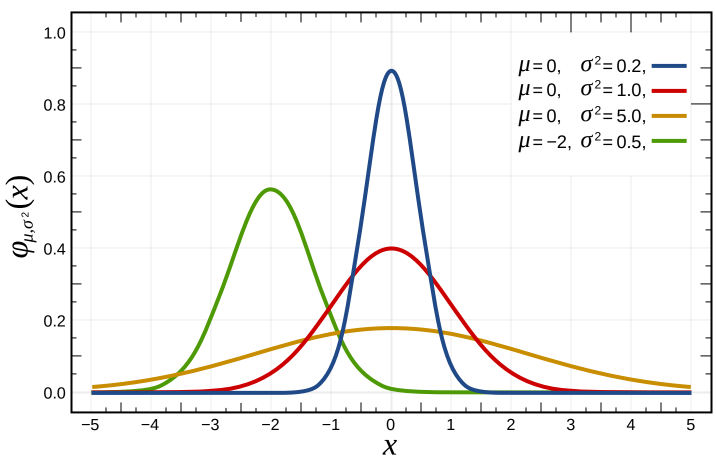

Figure 4: Gaussian distribution.[2]

(3) Exponential distribution: It is the probability distribution of the waiting time between events in a Poisson process in which events occur continuously and independently at a constant average

[1]Figure from Wikipedia: `https://en.wikipedia.org/wiki/Uniform_distribution_(continuous)`
[2]Figure from Wikipedia: `https://en.wikipedia.org/wiki/Normal_distribution`

rate (check Poisson process in later lectures)

$$f_X(x) = \begin{cases} \lambda e^{-\lambda} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{12}$$

The mean:

$$\begin{aligned}
\mathbb{E}[X] &= \int_{x=0}^{\infty} x f(x)\, dx \\
&= \int_{x=0}^{\infty} x \lambda \exp(-\lambda x)\, dx \\
&= \lambda \int_{x=0}^{\infty} x \exp(-\lambda x)\, dx \\
&= \lambda \left(\left[\frac{-1}{\lambda} x \exp(-\lambda x) \right]_{x=0}^{x=\infty} + \int_{x=0}^{\infty} \frac{1}{\lambda} \exp(-\lambda x)\, dx \right) \\
&= \lambda \left(0 + \frac{1}{\lambda^2} \right) \\
&= \frac{1}{\lambda}
\end{aligned}$$

Homework: *Find the variance of the exponential distribution.*

Answer:

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \int_{x=0}^{\infty} x^2 f(x)\, dx \\
&= \int_{x=0}^{\infty} x^2 \lambda \exp(-\lambda x)\, dx \\
&= \lambda \int_{x=0}^{\infty} x^2 \exp(-\lambda x)\, dx \\
&= \lambda \left(\left[\frac{-1}{\lambda} x^2 \exp(-\lambda x) \right]_{x=0}^{x=\infty} + \int_{x=0}^{\infty} \frac{1}{\lambda} x \exp(-\lambda x)\, dx \right) \\
&= \lambda \left(0 + \frac{1}{\lambda} \left(\frac{1}{\lambda} \mathbb{E}[X] \right) \right) \\
&= \lambda \left(\frac{1}{\lambda^3} \right) \\
&= \frac{1}{\lambda^2}
\end{aligned}$$

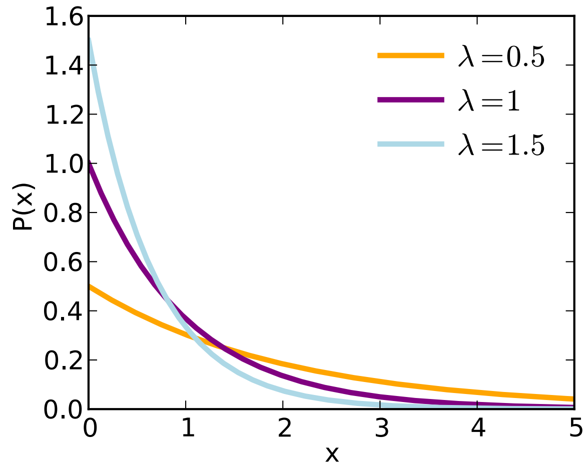

Figure 5: Exponential distribution.[3]

(4) Rayleigh Distribution:

$$f_X(x) \;=\; \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}}, x \geq 0, \tag{13}$$

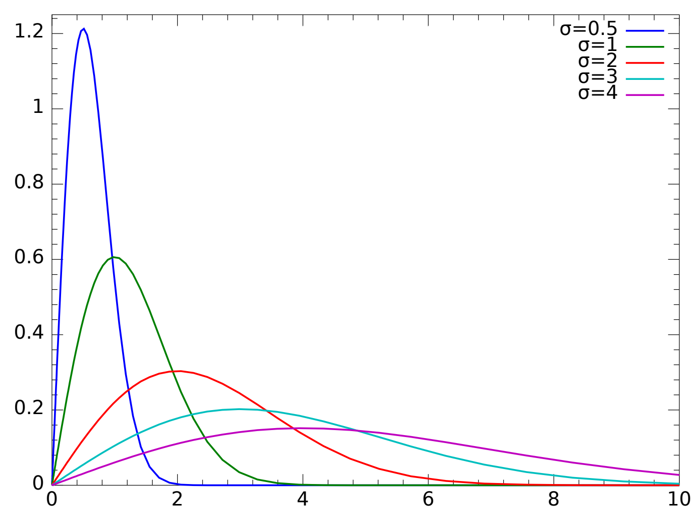

Figure 6: Rayleigh distribution.[4]

(5) Laplacian Distribution:

$$f_X(x) \;=\; \frac{1}{\sqrt{2}\sigma} e^{\frac{-\sqrt{2}|x|}{\sigma}}. \tag{14}$$

[3]Figure from Wikipedia: `https://en.wikipedia.org/wiki/Exponential_distribution`
[4]Figure from Wikipedia: `https://en.wikipedia.org/wiki/Rayleigh_distribution`

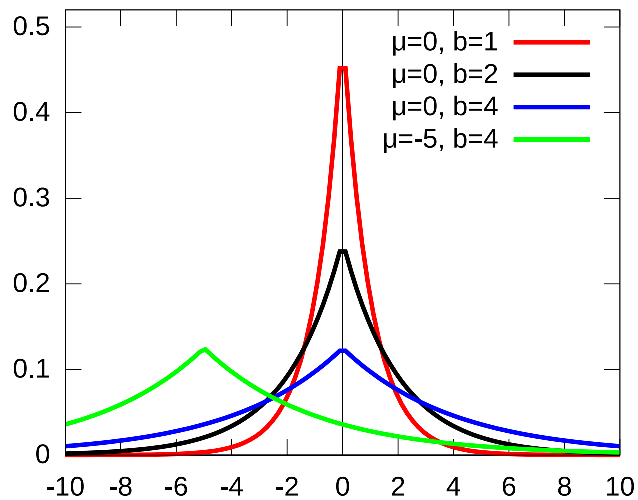

Figure 7: Laplacian distribution.[5]

1 Example of Discrete Random Variable

1.1 Bernoulli RV

flipping a coin, $P(H) = p$, $P(T) = 1 - p$, if head occurs $X = 1$, if tail occurs $X = 0$, $P(X = 0) = 1 - p, P(X = 1) = p$. The CDF of a bernoulli RV is as Figure 8.

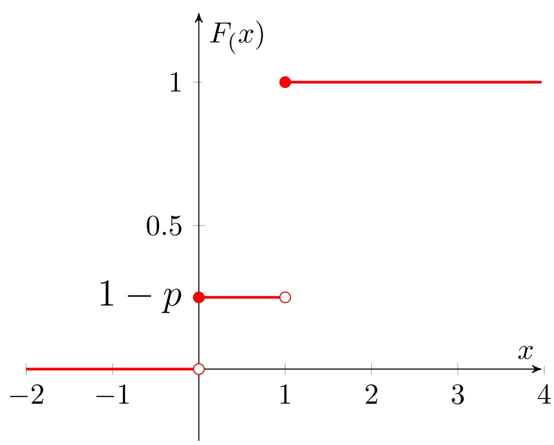

Figure 8: Cumulative distribution function of Bernoulli Random Variable

[5]Figure from Wikipedia: `https://en.wikipedia.org/wiki/Laplace_distribution`

1.2 Binomial distribution

Tossing a coin n times, $P(H) = p$, $P(T) = 1 - p$. X is number of heads, $x \in \{0, 1, \dots, n\}$.

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

Remark 2. *Let $Y_i \in \{0, 1\}$ denote the outcome of tossing the coin the ith time*

$$X = Y_1 + Y_2 + \cdots + Y_n.$$

i.e., a Binomial RV can be thought of as the sum of n independent Bernoulli RV.

Example 8 (Random graph)**.** *Each edge exists with probability p, X is the number of neighbor of node 1(Figure 9).*

$$Y_i = \begin{cases} 1, & \text{if node 1 is connected to i+1,} \\ 0, & \text{otherwise.} \end{cases}$$

$$X = Y_1 + Y_2 + \cdots + Y_{n-1}.$$

So X follows the Binomial distribution.

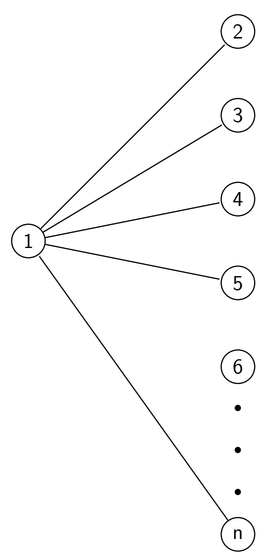

Figure 9: Random Graphs

Example 9 (BSC)**.** *Suppose we are transmitting a file of length n. Consider a BSC where the probability of error is p and the probability of receiving the correct bit is 1-p. (Figure 10) What is the probability that we have k errors?*

$$P(k\ errors) = \binom{n}{k} p^k (1-p)^{n-k}$$

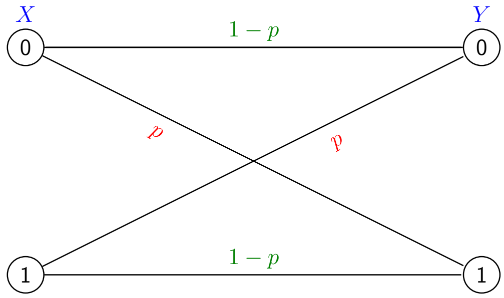

Figure 10: Binary Symmetric Channel with probability of error $P_e = p$.

Let X represent the number of errors, what is $E(X)$

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} kP(X=k), \\ &= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}, \\ &= np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}, \\ &= np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-k+1}, \\ &= np. \end{aligned}$$

Binomial theorem:

$$\begin{aligned} (x+y)^n &= \sum_{k=1}^{n}\binom{n}{k}x^k y^{n-k} \\ (p+1-p)^{n-1} &= \sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-k+1} \\ &= 1. \end{aligned}$$

Theorem 1. *For any two RVs X_1 and X_2, $Y = X_1 + X_2$,*

$$E(Y) = E(X_1) + E(X_2). \tag{15}$$

It does not matter whether X_1 and X_2 are independent or not.

1.3 Geometric distribution

You keep tossing a coin until you observe a Head. X is the number of times you have to toss the coin.

$$\begin{aligned} X &\in \{1, 2, \ldots\}, \\ P(X=K) &= (1-p)^k p. \end{aligned}$$

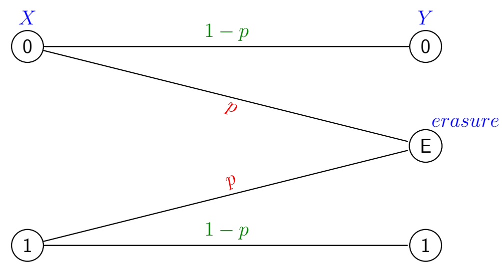

Figure 11: Binary Erasure Channel with probability of erasure $P_e = 0.1$.

Example 10 (Binary erasure channel)**.** *Suppose you have a BEC channel with feedback. When you get a erasure, you ask the sender to retransmit.(Figure 11)*
Suppose you pay one dollar for each retransmission. Let X be the amount of money you pay per transmission.

$$\begin{aligned} E(X) &= \frac{1}{1-p}, \\ &= \frac{1}{0.9} \approx 1.11\$. \end{aligned}$$

For geometric distribution,

$$P(H) \approx \frac{1}{E(X)},$$

which $E(X)$ is the number of coin flips on average.

Intuition: You have to make $E(X)$ trials, and in these $E(X)$ trials, the success happens once at the last trial

Proof.

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} kP(X=k), & (16) \\ &= \sum_{k=1}^{\infty} k(1-p)^{k-1}p, & (17) \\ &= p\sum_{k=1}^{\infty} k(1-p)^{k-1}. & (18) \end{aligned}$$

Recall that for $|x| < 1$,

$$\begin{aligned} \sum_{k=0}^{\infty} x^k &= \frac{1}{1-x}, & (19) \\ \frac{d}{dk}\sum_{k=1}^{\infty} kx^{k-1} &= \frac{1}{(1-x)^2}, & (20) \\ \sum_{k=1}^{\infty} k(1-p)^k &= \frac{1}{p^2}. & (21) \end{aligned}$$

So,

$$E(X) = p\frac{1}{p^2}, \tag{22}$$

$$= \frac{1}{p}. \tag{23}$$

□

1.4 Poisson distribution

Suppose a server receives λ searches per second on average. The probability that the server receives k searches for this second is

$$P(X = k) = C\frac{\lambda^k}{k!},\ k = 0, 1, 2, \ldots, \infty. \tag{24}$$

To find C:

$$\sum_{k=0}^{\infty} P(X = k) = 1$$
$$C\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1$$
$$C = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$
$$C = e^{-\lambda}.$$

Then the pdf of the poisson distribution for an average of λ arrivals per time unit is:

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!},\ k = 0, 1, 2, \ldots, \infty. \tag{25}$$

The mean is:

$$E(X) = \lambda.$$

Example 11 (Interpretation of poisson distribution as an arrival experiment)**.**
Suppose average of arrival cumtomers per second is λ. Suppose server goes down if $X \geq 100$. We want to find the probability of $P(X = k)$.

$$P(\textit{server going down}) = P(X \geq 100).$$

We divide the one second to n intervals, each length of the interval is $\frac{1}{n}$ second. The probability p of getting requests in small interval is $\frac{\lambda}{n}$.

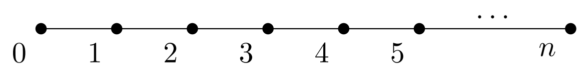

Figure 12: one second divided into n intervals.

Now we can consider it to be Bernoulli distribution with p.

$$\begin{aligned}
P(X=k) &= \binom{n}{k} p^k (1-p)^{n-k}, && (26)\\
&= \binom{n}{k} (\frac{\lambda}{n})^k (1-\frac{\lambda}{n})^{n-k} && (27)\\
&\approx \frac{n^k}{k!} (\frac{p}{1-p})^k (1-p)^n, && (28)\\
&= \frac{1}{k!} (np)^k e^{-np}, && (29)\\
&= \frac{1}{k!} \lambda^k e^{-\lambda}, && (30)\\
&= \frac{\lambda^k}{k!} e^{-\lambda}. && (31)
\end{aligned}$$

We get (28) *because of*

$$\begin{aligned}
\binom{n}{k} &= \frac{1}{k!} n(n-1)\dots(n-k+1), && (32)\\
&\approx \frac{n^k}{k!}, \ (k \text{ is a constant and } n \text{ goes to infinity}). && (33)
\end{aligned}$$

This means we can approximate Binomial(n,p) by Poisson with $\lambda = np$ (if n is very large).

2 Two Random Variables

Example 12. *Let X and Y be Bernoulli Random Variable. If $Y = 0$, we know X must equal to*

	Y=0	Y=1
X=0	$\frac{1}{2}$	$\frac{1}{4}$
X=1	0	$\frac{1}{4}$

Table 1: Joint probability mass function of X and Y.

0*, so X and Y are dependent.*

$$\begin{aligned}
P(X=0) &= \frac{3}{4},\\
P(X=1) &= \frac{1}{4}.
\end{aligned}$$

Here is a example which X and Y are independent, but they have the same marginal distribution.

	Y=0	Y=1
X=0	$\frac{3}{8}$	$\frac{3}{8}$
X=1	$\frac{1}{8}$	$\frac{1}{8}$

Table 2: Joint probability mass function of X and Y.

2.1 Marginalization

You have the joint distribution $P_{X,Y}(x,y)$.

$$P_X(x_0) = \sum_y P_{X,Y}(x_0,y), \quad (34)$$

$$P_Y(y_0) = \sum_x P_{X,Y}(x,y_0). \quad (35)$$

Definition 6. *If X and Y are continuous random variables, then the joint CDF:*

$$F_{X,Y}(x,y) = P(X \le x, Y \le y). \quad (36)$$

Given joint CDF $F_{X,Y}(x,y)$,

$$F_X(x_0) = F_{X,Y}(x_0,+\infty). \quad (37)$$

Definition 7. *When the CDF is differentiable, the joint pdf is defined as*

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}, \quad (38)$$

$$f_X(x) = \int_{\infty}^{-\infty} f_{X,Y}(x,y)dy, \quad (39)$$

$$f_Y(y) = \int_{\infty}^{-\infty} f_{X,Y}(x,y)dx. \quad (40)$$

Definition 8. *X and Y are independent if and only if*

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \quad (41)$$

$$f_{X,Y}(x,y) = f_X(x)f_Y(y). \quad (42)$$

Definition 9. *Conditional CDF of marginal distribution is*

$$F_{X,Y}(x|y) = P(X \le x | Y \le y), \quad (43)$$

$$= \frac{F_{X,Y}(X \le x, Y \le y)}{P(Y \le y)}. \quad (44)$$

Example 13. *X and Y are 2 random variables given by the joint pdf*

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} exp[\frac{-1}{2\sigma^2(1-\rho^2)}(x^2+y^2-2\rho xy)].$$

What is $f_X(x)$*?*

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy, \\
&= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} exp[\frac{-1}{2\sigma^2(1-\rho^2)}(x^2+y^2-2\rho xy)]dy, \\
&= \frac{exp[\frac{-x^2}{2\sigma^2(1-\rho^2)}]}{2\pi\sigma^2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} exp[\frac{-1}{2\sigma^2(1-\rho^2)}(y^2-2\rho xy+\rho x^2-\rho^2x^2)]dy, \\
&= \frac{exp[\frac{-x^2+\rho^2x^2}{2\sigma^2(1-\rho^2)}]}{2\pi\sigma^2\sqrt{\sigma\rho^2}} \int_{-\infty}^{\infty} e^{\frac{-(y-\rho x)}{2\sigma^2(1-\rho^2)}}dy, \\
&= \frac{exp[\frac{-(1-\rho^2)x^2}{2\sigma^2(1-\rho^2)}]}{\sqrt{2\pi}\sigma} \frac{\sqrt{1-\rho^{2^x}}}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\frac{-(y-\rho x)}{2\sigma^2(1-\rho^2)}}dy.
\end{aligned}$$

Because

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

So if $\rho = 0$,

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

Similarly,

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

We can have

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

So X *and* Y *are independent. If* $\rho \neq 0$, X *and* Y *are not independent.*