EE 381 Probability & Statistic with Applications to Computing (Fall 2020)

Lecture 8
Continuous Random Variables (p1)

Duc H. Tran, PhD

Continuos Random Variable’s topics

- **Probability distributions, Expectation**
- **Variance, Covariance & Correlation**
- Uniform Distribution, Normal Distribution
- Other Distributions, Central Limit Theorem

Continuous Random Variable

- **Example:** when measuring height, weight, and temperature, you have continuous data, because they can be divided into smaller increments, including fractional and decimal values.

Continuous Random Variable

- A non-discrete random variable X is said to be *absolute continuous*, or simply *continuous*, if its distribution function may be represented as:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \qquad (-\infty < x < \infty)$$

Where the function *f(x)* has the properties

❖ $f(x) \geq 0$

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

- The probability that X takes on any one particular value is zero.
- The *interval probability* that X lies between two different values, a and b, is given by:

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$

Continuous Random Variable

- **Example:** Assuming we have function *f(x)* as follow:

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & otherwise \end{cases}$$

a) Find the constant *c* such that *f(x)* is a density function.

b) Compute *P(1 < x < 2)*.

Solution:

a) f(x) must satisfy two properties: $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$

Satisfying property 1 → $c \geq 0$

Satisfying property 2 → $\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{3} cx^2\, dx = \left.\frac{cx^3}{3}\right|_{0}^{3} = 9c = 1$ → c=1/9.

b) $P(1 < X < 2) = \int_{1}^{2} \frac{1}{9} x^2 dx = \left.\frac{x^3}{27}\right|_{1}^{2} = \frac{7}{27}$

Continuous Random Variable

- **Example (ctn.):** Assuming we have function *f(x)* as follow:

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & otherwise \end{cases}$$

c) Find the distribution function F(x).

d) Compute $P(1 < x \leq 2)$.

Solution:

c) $\mathrm{F(x)} = \mathrm{P(X} \leq x) = \int_{-\infty}^{x} f(u)du$

If x<0, then F(x)=0. If $0 \leq x < 3$, then:

$$F(x) = \int_0^x f(u)du = \int_0^x \frac{1}{9}u^2 du = \frac{x^3}{27}$$

If $x \geq 3$, then

$$F(x) = \int_0^3 f(u)du + \int_3^x f(u)du = \int_0^3 \frac{1}{9}u^2 du + \int_3^x 0du = 1$$

The required distribution function F(x) is:

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^3}{27} & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Continuous Random Variable

d) Compute $P(1 < x \leq 2)$.

Solution:

d) We have:

$$\begin{aligned} P(1 < X \leq 2) &= P(X \leq 2) - P(X \leq 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

Using this result, we can easily prove that:

$$P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2) = \frac{7}{27}$$

In-class Exercise

Let X is random variable which has the following density function *f(x)*:

$$f(x) = \begin{cases} \frac{x}{2} & 0 \le x \le 2 \\ 0 & otherwise \end{cases}$$

1) Find $P(\frac{1}{2} < X < \frac{3}{2})$

a) 1/2 b) 1/3

c) 1/4 d) 1/5

2) Find the distribution function

a) $F(x) = \begin{cases} 0 & x \le 0 \\ \frac{3x^2}{4} & 0 \le x \le 2 \\ 1 & x \ge 2 \end{cases}$

b) $F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^2}{2} & 0 \le x \le 2 \\ 1 & x \ge 2 \end{cases}$

c) $F(x) = \begin{cases} 0 & x \le 0 \\ \frac{5x^2}{4} & 0 \le x \le 2 \\ 1 & x \ge 2 \end{cases}$

d) $F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^2}{4} & 0 \le x \le 2 \\ 1 & x \ge 2 \end{cases}$

Continuous Random Variable

The probability that X is between x and $x + \Delta x$ is given by:

$$P(x \leq X \leq x + \Delta x) = \int_{x}^{x+\Delta x} f(u)\, du$$

So that if Δx is small, we have: $P(x \leq X \leq x + \Delta x) = f(x)\Delta x$

It is also observed that when f(x) is continuous, the derivative of the distribution function is the density function:

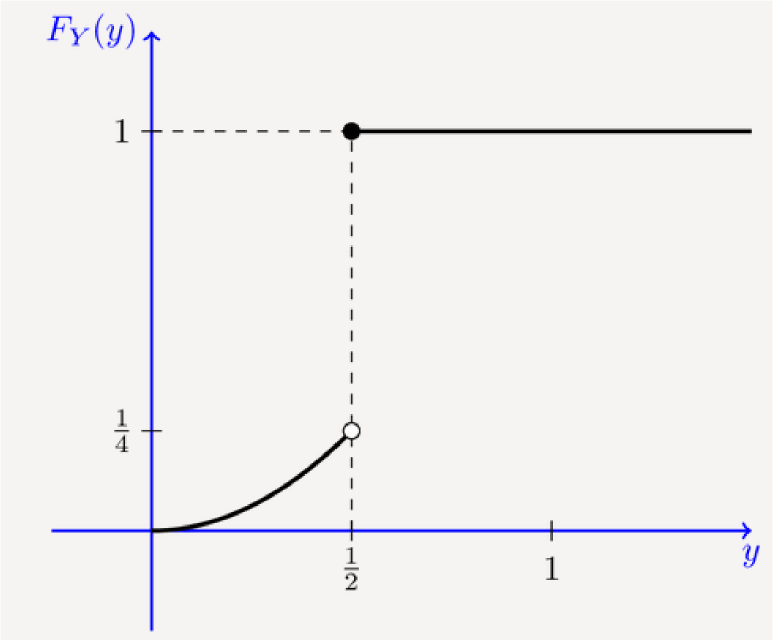
$$\frac{dF(x)}{dx} = f(x)$$

Continuous Random Variable

Note that there are also random variables exist that are neither discrete nor continuous.

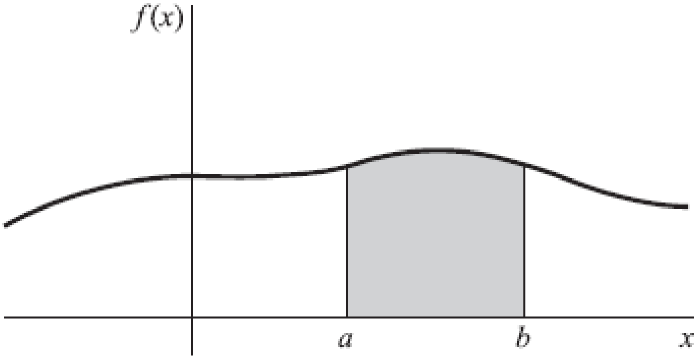
We called them mixed random variables, which contain a discrete part and a continuous part.

$$F_Y(y) = \begin{cases} 1 & y \geq \frac{1}{2} \\ y^2 & 0 \leq y < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$



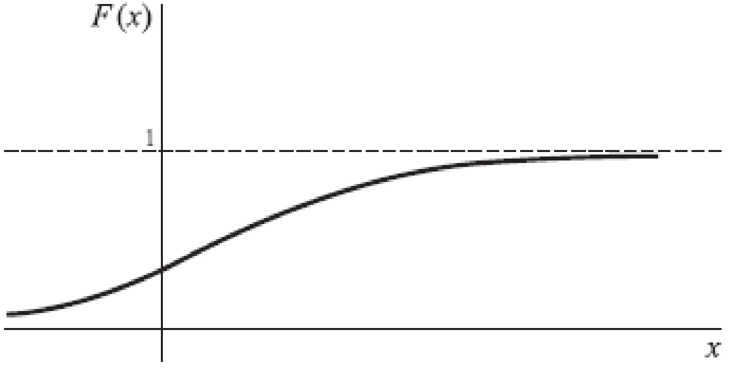
Graphical Interpretations

- If *f(x)* is the density function for a random variable *X*, then we can represent *y=f(x)* graphically by a curve as follow:



$f(x) \geq 0$ so the curve cannot fall below the *x* axis. *P(a<X<b)* is the shaded area.

- The distribution function is a monotonically increasing function which increases from 0 to 1 and is represented by a curve as follow:



Joint Distribution for Continuous Random Variables

- The joint probability function for random variables X and Y is defined by:
  1. $f(x, y) \geq 0$
  2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$
- The probability that X lies btw a&b while Y lies btw c&d is given by:

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b} \int_{y=c}^{d} f(x, y) dx dy$$

- The joint distribution function of X and Y is defined by:

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{y} f(u, v)\, du dv$$

- Then the density function is obtained by differentiating the distribution function with respect to x and y:

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

Joint Distribution for Continuous Random Variables

- The distributions functions of X and Y are defined by:

$$P(X \leq x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u,v)dudv$$

$$P(Y \leq y) = F_2(y) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{y} f(u,v)dudv$$

The density functions of X and Y are defined by:

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x,v)dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u,y)du$$

In-class Exercise

Suppose that random variables X and Y have a joint density function given by

$$f(x,y) = \begin{cases} \dfrac{2x+y}{210} & 2 < x < 6, 0 < y < 5 \\ 0 & otherwise \end{cases}$$

1) Find P$(3 < x < 4, y > 2)$. (Choose the correct answer below)
   a) 1/20 b) 3/20
   c) 5/20 d) 7/20

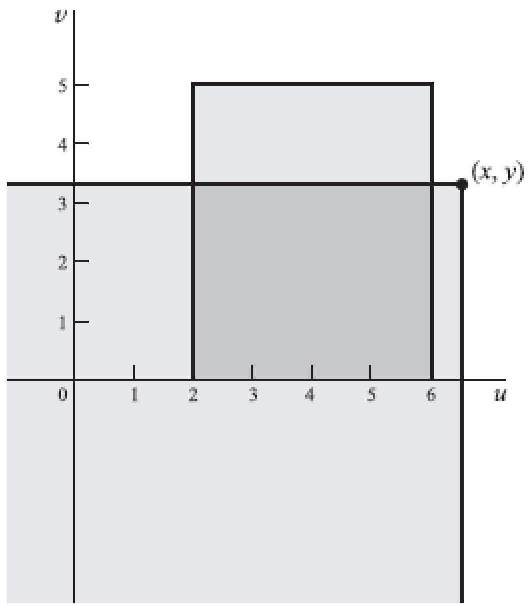
2) Find the joint distribution function of X and Y when x>6 and 0<y<5. (Choose the correct answer below)
   a) $\frac{16y+y^2}{105}$ b) $\frac{17y+y^2}{105}$
   c) $\frac{18y+y^2}{105}$ d) $\frac{19y+y^2}{105}$

In-class Exercise

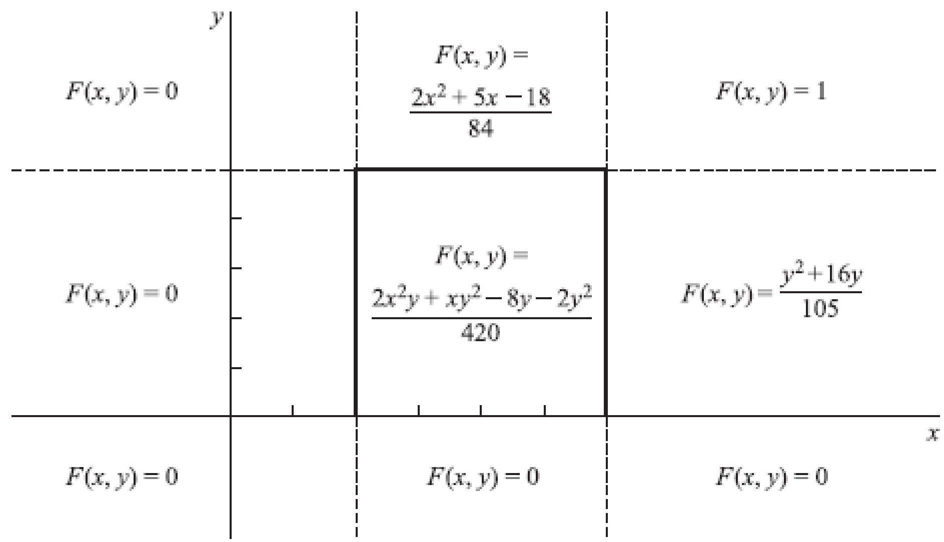
Suppose that random variables X and Y have a joint density function given by

$$f(x,y) = \begin{cases} \dfrac{2x+y}{210} & 2 < x < 6, 0 < y < 5 \\ 0 & otherwise \end{cases}$$

Question 2's explanation: Find the joint distribution function of X and Y



Question 2



Complete F(x, y)

Changes of Variables

- **Theorem 2-3**: Let X be a continuous random variable with probability density f(x). Let us define $U = \phi(X)$, where to each value of X there corresponds one and only one value of U and conversely, so that $\mathrm{X} = \psi(U)$. Then the probability density of U is given by *g(u)* where

$$g(u)|du| = f(x)|\mathrm{dx}|$$

or

$$g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(U)]|\psi'(u)|$$

Changes of Variables

Example: The probability function of random variable X is given by:

$$f(x) = \begin{cases} \dfrac{x^2}{81} & -3 < x < 6 \\ 0 & otherwise \end{cases}$$

Find the probability density for the random variable $U = \frac{1}{3}(12 - X)$.

Solution: $x = 12 - 3u$. So $u = 5$ when $x = -3$, and $u = 2$ when $x = 6$.

Since $\psi'(u) = \frac{dx}{du} = -3$, follows Theorem 2-3, the density function for U is:

$$g(u) = \begin{cases} \dfrac{(12 - 3u)^2}{27} & 2 < u < 5 \\ 0 & otherwise \end{cases}$$

Recheck: $\int_2^5 \frac{(12-3u)^2}{27} du = -\frac{(12-3u)^3}{243} \Big|_2^5 = 1$

Changes of Variable

Theorem 2-4: Let X and Y be a continuous random variables having joint probability function *f(x, y).* Let's define $U = \phi_1(X,Y), V = \phi_2(X,Y)$, where to each pair of X and Y there corresponds one and only one pair of values of U and V conversely, so that $X = \psi_1(U,V), Y = \psi_2(U,V)$. Then the joint probability function for U and V is given by *g(u,v)* where

$$g(u,v)|du\ dv| = f(x,y)|\mathrm{dx\ dy}|$$

Or
$$g(u,v) = f(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = f[\psi_1(u,v),\psi_2(u,v)]|J|$$

The Jacobian determinant, or briefly Jacobian, is given by:

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

Changes of Variables

Example: If the random variables X and Y have joint density function:

$$f(x, y) = \begin{cases} \dfrac{xy}{96} & 0 < x < 4, 1 < y < 5 \\ 0 & otherwise \end{cases}$$

Find the probability density for the random variable $U = (X + 2Y)$.

Solution: Let u $= x + 2y$, $v = x$ (chosen arbitrarily). This yields $x = v$, $y = \frac{1}{2}(u\text{-}v)$

$0 < x < 4, 1 < y < 5$ corresponds to $0 < v < 4,\ \ 2 < u - v < 10.$

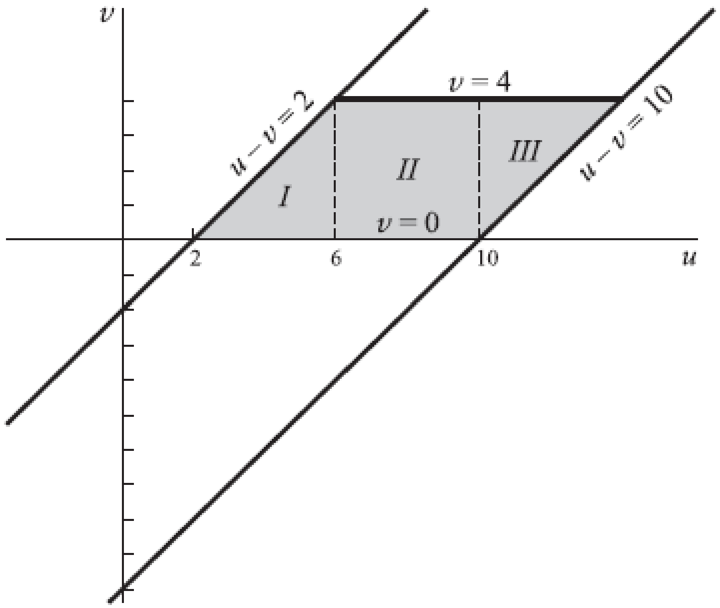
The Jacobian is given by: $J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$

By Theorem 2-4, the joint density function of U and V is:

$$g(u, v) = \begin{cases} \dfrac{v(u - v)}{384} & 2 < u - v < 10, 0 < v < 4 \\ 0 & otherwise \end{cases}$$

Changes of Variables

Example (ctn.): The marginal density function of U is given by:



$$g_1(u) = \begin{cases} \displaystyle\int_{v=0}^{u-2} \frac{v(u-v)}{384}\,dv & 2 < u < 6 \\ \displaystyle\int_{v=0}^{4} \frac{v(u-v)}{384}\,dv & 6 < u < 10 \\ \displaystyle\int_{v=u-10}^{4} \frac{v(u-v)}{384}\,dv & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

Carrying out the integrations:

$$g_1(u) = \begin{cases} (u-2)^2(u+4)/2304 & 2 < u < 6 \\ (3u-8)/144 & 6 < u < 10 \\ (348u - u^3 - 2128)/2304 & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

In-class Exercise

Suppose that random variables X and Y have a joint density function given by

$$f(x,y) = \begin{cases} \dfrac{2x+y}{210} & 2 < x < 6, 0 < y < 5 \\ 0 & otherwise \end{cases}$$

Find the joint density function of $U = (X + Y)$ and $V = X$.

Choose the correct answer:

a) $g(u,v) = \begin{cases} \frac{u+2v}{210} & 0 < u - v < 5, 2 < v < 6 \\ 0 & otherwise \end{cases}$

b) $g(u,v) = \begin{cases} \frac{2u+v}{210} & 0 < u - v < 5, 2 < v < 6 \\ 0 & otherwise \end{cases}$

c) $g(u,v) = \begin{cases} \frac{u+v}{210} & 0 < u - v < 5, 2 < v < 6 \\ 0 & otherwise \end{cases}$

d) $g(u,v) = \begin{cases} \frac{u-v}{210} & 0 < u - v < 5, 2 < v < 6 \\ 0 & otherwise \end{cases}$

Conditional Distributions

- Let X and Y are continuous random variables. The conditional density function of Y given X is:

$$f(y|x) = \frac{f(x,y)}{f_1(x)}$$

Where *f(x,y)* is the joint density function of X and Y; $f_1(x)$ is the marginal density function of X.

- Example: Find the probability of Y being between c and d given that $x < X < x+dx$.

$$P(c < Y < d | x < X < x + dx) = \int_c^d f(y|x)\, dy$$

Conditional Distributions

Example: if X and Y have the joint density function

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

- Find (a) $f(y|x)$, (b) $P(Y > \frac{1}{2} | \frac{1}{2} < X < \frac{1}{2} + dx)$.

Solution:

a) For 0 < x < 1, we have $$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{other } y \end{cases}$$

b) $$P(Y > \tfrac{1}{2} \,|\, \tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y \,|\, \tfrac{1}{2})\, dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\, dy = \frac{9}{16}$$

Expectation, Variance, Covariance & Correlation

- Expectation of discrete random variables: $E(X) = \mu_X = \sum xf(x)$
- Expectation of continuous random variables: $E(X) = \mu_X = \int_{-\infty}^{\infty} xf(x)dx$

Example: The density function of a random variable X is given by

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

The expected value of X is:

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^2 x\left(\frac{1}{2}x\right)dx = \int_0^2 \frac{x^2}{2}\,dx = \frac{x^3}{6}\bigg|_0^2 = \frac{4}{3}$$

Expectation, Variance, Covariance & Correlation

- *Variance* of continuous random variables:

$$\sigma_X^2 = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

Standard deviation: $\sigma = \sqrt{\sigma^2}$

Example: The density function of a random variable *X* is given by

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

The variance of *X* is ($\mu = 4/3$):

$$\sigma^2 = E\left[\left(X - \frac{4}{3}\right)^2\right] = \int_{-\infty}^{\infty}\left(x - \frac{4}{3}\right)^2 f(x)\,dx = \int_0^2\left(x - \frac{4}{3}\right)^2\left(\frac{1}{2}x\right)dx = \frac{2}{9}$$

Expectation, Variance, Covariance & Correlation

Moment generating function of random variable X:

$$M_X(t) = \sum e^{tx} f(x) \qquad \text{(discrete variable)}$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \qquad \text{(continuous variable)}$$

Expectation, Variance, Covariance & Correlation

*Covariance (*variance for joint distributions*)*:

If X and Y are 2 continuous random variables having joint density function f(x,y). The means of X and Y are:

$$\mu_X = E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy, \qquad \mu_Y = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy$$

And the variance are:

$$\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)^2 f(x, y)\,dx\,dy$$

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - \mu_Y)^2 f(x, y)\,dx\,dy$$

The covariance is defined as:

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y)\,dx\,dy$$

Expectation, Variance, Covariance & Correlation

Correlation for continuous random variable has the same formula as correlation for discrete random variables.

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Reference

Notes, equations, and figures in the lecture are based on or taken from materials in the course textbook:

"Probability and Statistics", by Spiegel, Schiller and Srinivasan, ISBN 987-007-179557-9 (McGraw-Hill/Schaun's)