Random Variables and Probability Distributions.

Example: $S = \{HH, HT, TH, TT\}$

Sample Point	HH	HT	TH	TT
X	2	1	1	0

Discrete Probability Distributions: Let X random variable and outcomes $x_1, x_2, x_3, \ldots$

Suppose $P(X = x_k) = f(x_k)$, $k = 1, 2, \ldots$

The probability distribution is $P(X = x) = f(x)$.

Rules

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$.

Example:

$P(X=0) = P(TT) = 1/4$.

$P(X=1) = P(HT \cup TH) = P(HT) + P(TH) = 1/2$.

$P(X=2) = P(HH) = 1/4$

Probability function is

x	0	1	2
$f(x)$	1/4	1/2	1/4.

Distribution Functions: Cumulative distribution function for random variable X is $F(x) = P(X \leq x)$, $-\infty < x < \infty$.

1. $F(x)$ is nondecreasing i.e $F(x) \leq F(y)$ if $x \leq y$.
2. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$
3. $F(x)$ is continuous from the right i.e. $\lim_{h \to 0^+} F(x+h) = F(x)$.

Distribution function for discrete random variable X, for all $x \in (-\infty, \infty)$:

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u).$$

if X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, the distribution function

$$F(x) = \begin{cases} 0 & -\infty < x < \infty \\ f(x_1) & x_1 \leq x \leq x_2 \\ f(x_1) + f(x_2) & x_2 \leq x \leq x_3 \\ \vdots & \vdots \\ f(x_1) + \ldots + f(x_n) & x_n \leq x \leq \infty . \end{cases}$$

Example. Using coin-toss example :

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/4 & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty . \end{cases}$$

Continuous Random Variables: X is **continuous** if distribution function :

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

Interval Probability: $$P(a < X < b) = \int_a^b f(x)\,dx.$$

Example : Find **c** such that following function is (a) density function (b). $P(1 < X < 2)$:

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

(a). $$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^3 cx^2\,dx = \left.\frac{cx^3}{3}\right|_0^3 = 9c. \quad \therefore c = 1/9.$$

(b). $$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}.$$

note as $f(x)$ is continuous : $P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2)$.

Example : Find the distribution function and use result to find $P(1 < x \leq 2)$

(a) $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$.

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

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

If $x \geqslant 3$,

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

(b). $P(1 < X \leq 2) = P(X \leq 2) - P(X \leq 1) = F(2) - F(1) = 2^3/27 + 1^3/27 = 7/27.$

Note: Probability X is between x and $x + \Delta x$ is :

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

If $\Delta x \ll 1$,

$P(x \leq X \leq x + \Delta x) = f(x)\Delta x$, and differentiating both sides :

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

Discrete case: The joint probability function X and Y is :

$P(X = x, Y = y) = f(x, y)$ where

1. $f(x,y) \geqslant 0$
2. $\sum_x \sum_y f(x,y) = 1$. (sum of all values of x and y is 1)

Example: Find the distribution function and use result to find $P(1 < x \leq 2)$

(a). $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$.

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

$$F(x) = \int_{0}^{x} f(u)\,du = \int_{0}^{x} \frac{1}{9}u^2\,du = x^3/27.$$

If $x \geqslant 3$,

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

(b). $P(1 < X \leq 2) = P(X \leq 2) - P(X \leq 1) = F(2) - F(1) = 2^3/27 + 1^3/27 = 7/27.$

Note: Probability X is between x and $x + \Delta x$ is:

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

If $\Delta x \ll 1$,

$P(x \leq X \leq x + \Delta x) = f(x)\Delta x$, and differentiating both sides:

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

Discrete case: The joint probability function X and Y is:

$P(X = x, Y = y) = f(x, y)$ where

1. $f(x,y) \geqslant 0$
2. $\sum_{x}\sum_{y} f(x,y) = 1$. (sum of all values of x and y is 1)

Probability that $X = x_j$ is obtained by adding all entries in row given by:

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^{n} f(x_j, y_k).$$

Similarly for $Y = y_k$:

$$P(Y = y_j) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k).$$

Total probability: $\sum_{j=1}^{m} \sum_{k=1}^{n} f(x_j, y_k) = 1.$

Joint Distribution: $F(x,y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(u,v).$

Continuous Case: The joint density function of X and Y is:

1. $f(x,y) \geq 0$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1.$

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

Generally if A is any event there will be a region R_A s.t.

$$P(A) = \iint_{\mathcal{R}} f(x,y)\, dx\, dy.$$

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

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

Density Function: $P(X \leq x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u,v)\, du\, dv.$

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

Independent Random Variables: Suppose X and Y are random variables, if events $X=x$, $Y=y$ are independent:

$$P(X=x, Y=y) = P(X=x)P(Y=y), \quad \text{or} \quad f(x,y) = f_1(x)f_2(y).$$

Dependent: If $f(x,y)$ cannot be expressed then X and Y are dependent.

They are independent if $X \leq x$ and $Y \leq y$ are independent events for all x and y:

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y).$$

Change of variables:

1. X discrete random variable with probability $f(x)$

   Suppose U is defined in terms of X by $U = \phi(X)$ where each value of U corresponds to only one value of U, s.t. $X = \psi(U)$.

   Probability function of U is:

   $$g(u) = f[\psi(u)].$$

2. X, Y discrete random variables with joint probability $f(x,y)$.

   $U = \phi_1(X,Y)$ and $V = \phi_2(X,Y)$

   $X = \psi_1(U,V)$ and $Y = \psi_2(U,V)$

   Joint probability is:

   $$g(u,v) = f[\psi_1(u,v), \psi_2(u,v)].$$

Continuous Variables:

1. X continuous random variable, $U = \phi(X)$ and $X = \psi(U)$, probability density:

   $$g(u)\,|du| = f(x)\,|dx|$$

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

2. X, Y joint density $f(x,y)$, $U = \phi_1(X,Y)$ and $V = \phi_2(X,Y)$

   Let $X = \psi_1(U,V)$ and $Y = \psi_2(U,V)$.

   Joint density function:

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

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

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

Theorems:

1. X, Y continuous random variables, $U = \phi_1(X,Y)$, $V = X$ (2nd choice arbitrary)
   Joint density U is marginal density obtained from joint density of U and V.

2. Let $f(x,y)$ be joint density function of X and Y.
   Then $g(u)$ density function of random variable $U = \phi_1(X,Y)$ is found by differentiating wrt u the distribution function:

$$G(u) = P[\phi_1(X,Y) \leq u] = \iint_{\mathcal{R}} f(x,y)\,dx\,dy.$$

where $\mathcal{R}$ is region for which $\phi_1(x,y) \leq u$

Convolutions: Density function of two continuous random variables X and Y is:

$$g(u) = \int_{-\infty}^{\infty} f(x, u-x)\,dx$$

In special case where X and Y are independent, $f(x,y) = f_1(x) f_2(y)$ reduces:

$$g(u) = \int_{-\infty}^{\infty} f_1(x) f_2(x-u)\,dx \quad (\text{convolution, } f_1 * f_2).$$

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 * (f_2 + f_3) = f_1 * f_2 + f_1 * f_3$.

Conditional Distributions:

If $P(A) > 0$, $P(B|A) = P(A \cap B)/P(A)$

If X, Y discrete random variables $(A : X = x)$, $(B : Y = y)$ then:

$$P(Y=y \mid X=x) = f(x,y) / f_1(x) \quad (P(X=x, Y=y) \text{ joint probability}).$$

$f_1(x)$ is the marginal probability function for X.

$$f(y|x) \equiv f(x,y) / f_1(x),$$ similarly:

$$f(x|y) \equiv f(x,y) / f_2(y)$$

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