Joint Distribution

Similar to the one-dimensional situation, we can denote the range space of (X, Y) by:

$$R\_\{X, Y\} = \{(x, y) \mid x = X(s), y = Y(s), s \in S\}$$

$(X(s), Y(s))$ is a discrete two-dimensional random variable if the number of possible values of $(X(s), Y(s))$ are finite or countable. That is, the possible values of $(X(s), Y(s))$ may be represented by $(X(s), Y(s)), i = 1,2,3, \dots; j = 1,2,3, \dots$

(X, Y) is a continuous two-dimensional random variable if the possible values of $(X(s), Y(s))$ can assume any value in some region of the Euclidean space R^2

The joint probability (mass) function of a discrete random variable is

$f_{X,Y}(x, y) = P(X = x, Y = y)$ and

$f_{X,Y}(x, y) \geq 0$ for x,y in the range of x,y

$f_{X,Y}(x, y) = 0$ for x,y not in the range of x,y

$$\sum \sum_{(x,y) \in R_{X,Y}} f_{X,Y}(x, y) = 1$$

$$\sum \sum_{(x,y) \in A} f_{X,Y}(x, y) = P((X, Y) \in A)$$

The joint probability density function of a continuous random variable is

$$P((X, Y) \in D) = \int \int_{(x,y) \in D} f_{X,Y}(x, y) \; dy \, dx$$

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x, y) \; dy \, dx$$

Marginal Probability Distribution of x

Discrete:

$$f_X(x) = \sum_y f_{X,Y}(x, y)$$

Continuous:

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

Conditional Probability Distribution of Y given X = x:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

(the distribution of Y given that the random variable X is observed to take the value x)

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

(the distribution of X given that the random variable Y is observed to take the value y)

$$P(Y \leq y | X = x) = \int_{-\infty}^{y} f_{Y|X}(y|x) \, dy$$

$$E(Y | X = x) = \int_{-\infty}^{\infty} y \, f_{Y|X}(y|x) \, dy$$

Random variables X and Y are independent if and only if for any x and y,

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

Properties of Independent Random Variables

If X and Y are independent random variables, the following properties hold:

1. For any arbitrary subsets A and B of R, the events X ∈ A and Y ∈ B are independent events in S. Thus, $P(X \in A; Y \in B) = P(X \in A)P(Y \in B)$. In particular, for any real numbers x, and y, we have $P(X \leq x; Y \leq y) = P(X \leq x)P(Y \leq y)$.
2. Independence is connected with conditional distribution

Expectation

Consider a 2 variable function $g(x, y)$:

Remember $E(g(X)) = \sum g(x) f(x) \; or \int_{-\infty}^{\infty} g(x) f(x) dx$If (X, Y) is a discrete random variable,

$$E(g(X, Y)) = \sum_x \sum_y g(x, y) f_{X,Y}(x, y)$$

If (X, Y) is a continuous random variable,

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) \; dy \, dx$$

If \$$g(X, Y) = (X - E(X))(Y - E(Y)) = (X - \mu_X)(Y - \mu_y)$, the expectation $E[g(x, y)]$ leads to the covariance of X and Y.

Covariance

$$cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

If $(X, Y)$is a discrete random variable,

$$cov(X, Y) = \sum_x \sum_y (x - \mu_x)(y - \mu_y) f_{X,Y}(x, y)$$

If $(X, Y)$is a continuous random variable,

$$cov(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f_{X,Y}(x, y) \; dy \, dx$$

$$cov(X, Y) = cov(Y, X);$$

$$cov(X + b, y) = cov(X, Y);$$

$$cov(aX, Y) = a \times cov(X, Y)$$

$$V(aX + bY) = a^2 V(X) + b^2 V(Y) - 2ab \times cov(X, Y)$$

Random Variables

Probability mass function

For a discrete random variable X, define

$f(x) = P(X = x) \; for \; all \; x \in R_x \, , 0 \; for \; all \; x \notin R_x$n , $f(x)$ is the probability function or probability mass function.

$$f(x_i) \geq 0 \; for \; all \; x_i \in R_X$$
$$f(x) = 0 \; for \; all \; x \notin R_X$$
$$\sum_{i=1}^{\infty} f(x_i) = 1, or \sum_{x_i \in R_X} f(x_i) = 1$$
$$For \; any \; set \; B \subset R, we \; have \; P(X \in B) = \sum_{x_i \in B \cap R_X} f(x_i).$$

Probability density function

For a continuous random variable X, define

$$f(x) \geq 0 \; for \; x \in R_x \, , f(x) = 0 \; for \; x \notin R_x$$

$$\int_{R_x} f(x) dx = 1$$

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$

Notice that

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

$$P(X = x_0) = 0$$

$$P(a \leq x \leq b) = P(a < x < b) = P(a \leq x < b) = P(a < x \leq b)$$

Cumulative distribution function

For a random variable X,

$$F(x) = P(X \leq x)$$

CDF of discrete random variable

$$F(x) = \sum_{t \in R_X; t \leq x} P(X = t)$$

$$P(a \leq x \leq b) = F(b) - F(a-)$$

a- is the largest R_x smaller than a. Notice that $P(x < a) = F(a-)$

CDF of continuous random variable

$$F_x(x) = \int_{-\infty}^{x} f(t)dt$$

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

$$P(a \leq x \leq b) = P(a < x < b) = F(b) - F(a)$$

Expectation

For a discrete random variable:

$$E(X) = \mu_x = \sum_{x_i \in R_X} x_i\, f(x_i)$$

For a continuous random variable: $E(X) = \int_{x \in R_X} x f(x) dx$

Properties:

$$E(aX + b) = aE(X) + b$$

$$E(X + Y) = E(X) + E(Y)$$

Variance

$$\sigma_x^2 = V(X) = E(X - \mu_x)^2 = \sum_{x \in R_X} (x - \mu_x)^2 f(x) \text{ } or \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x)$$

Properties: $V(aX + b) = a^2 V(X)$ $V(X) = E(X^2) - [E(X)]^2$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(X + Y) = V(X) + V(Y) + 2cov(X, Y)$$

Basic Concepts of Probability

A **statistical experiment** is any procedure that produces data or observations. The **sample space**, denoted by S, is the set of all possible outcomes of a statistical experiment. The sample space depends on the problem of interest! A **sample point** is an outcome (element) in the sample space. An **event** is a subset of the sample space.

Conditional Probability

A statistical experiment is any procedure that produces data or observations. The sample space, denoted by S, is the set of all possible outcomes of a statistical experiment. The sample space depends on the problem of interest! A sample point is an outcome (element) in the sample space. An event is a subset of the sample space.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Multiplication rule

$$P(A \cap B) = P(B|A)P(A)$$

Inversion probability rule

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

Independence

If A and B are independent:

$$P(A \cap B) = P(A)P(B)$$

$$P(B|A) = P(B)$$

Law of total probability

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

For a special case with any events A and B: $P(B) = P(A)P(B|A) + P(A')P(B|A')P(B) = P(A \cap B) + P(A' \cap B)$

Bayes Theorem

We can also see, as an extension of the inversion probability rule

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

Things to note:

$$P(B \cap A) = P(A) - P(B' \cap A)$$

$$P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$$