CHAPTER 3 - PROBABILITY DISTRIBUTIONS AND PROBABILITY DENSITIES

3.1 Random Variables

Definition 3.1 (Random Variable)**.** If S is a sample space with a probability measure and X is a real-valued function defined over the elements of S, then X is called a random variable.

Discrete Random Variables - random variables whose range is finite or countably infinite

3.2 Probability Distributions

Definition 3.2 (Probability Distribution)**.** If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range of X is called the probability distribution of X.

Theorem 3.1 (Probability Distribution - Discrete Conditions)**.** *A function can serve as the probability distribution of a discrete random variable X if and only if its values, $f(x)$, satisfy the conditions:*

(1) *$f(x) \geq 0$ for each value within its domain*
(2) *$\sum_x f(x) = 1$, where the summation extends over all the values within its domain*

Definition 3.3 (Distribution Function)**.** If X is a discrete random variable, the function given by

$$F(x) = P(X \leq x) = \sum_{t \leq X} f(t) \qquad \text{for} -\infty < x < \infty$$

where $f(t)$ is the value of the probability distribution of X at t, is called the **distribution function**, or the **cumulative distribution** of X

Theorem 3.2 (Distribution Function - Discrete Conditions)**.** *The values $F(x)$ of the distribution function of a discrete random variable X satisfy the conditions:*

(1) *$F(-\infty) = 0$ and $F(\infty) = 1$*
(2) *if $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b*

Proof. TBD: Exercise 3.8 □

Theorem 3.3 (Probability Distribution from Cumulative Distribution - Discrete)**.** *If the range of a random variable X consists of the values $x_1 < x_2 < x_3 < \cdots < x_n$, then $f(x_1) = F(x_1)$ and*

$$f(x_i) = F(x_i) - F(x_{i-1}) \qquad \textit{for } i = 2, 3, \ldots, n$$

3.3 Continuous Random Variables

This section intentionally left blank

3.4 Probability Density Functions

Definition 3.4 (Probability Density Function)**.** A function with values $f(x)$, defined over the set of all real numbers, is called a **probability density function** or abreviated as **pdf** of the continuous random variable X if and only if

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

for any real constants a and b with $a \leq b$

Theorem 3.4 (Endpoints not needed)**.** *If X is a continuous random variable and a and b are real constants with $a \leq b$, then*

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

Theorem 3.5 (Probability Density - Continuous Conditions)**.** *A function can serve as a probability density of a continuous random variable X if its values, $f(x)$, satisfy the conditions:*

(1) $f(x) \geq 0$ *for* $-\infty < x < \infty$
(2) $\int_{-\infty}^{\infty} f(x)dx = 1$

Definition 3.5 (Distribution Function)**.** If X is a continuous random variable and the value of its probability density at t is $f(t)$, then the function given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \qquad \text{for } -\infty < x < \infty$$

is called the **distribution function** or the **cumulative distribution function** of X

Theorem 3.6 (Derivative of Distribution Function is Probability Density)**.** *If $f(x)$ and $F(x)$ are the values of the probability density and the distribution function of X at x, then*

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

for any real constants a and b with $a \leq b$, and

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

where the derivative exists.

3.5 Multivariate Distributions

bivariate - situations where we are interested at the same time in a pair of random variables defined over a joint sample space **multivariate** - situations covering any finite number of random variables **univariate** - situations with one random variable

Definition 3.6 (Joint Probability Distribution)**.** If X and Y are discrete random variables, the function given by $f(x, y) = P(X = x, Y = y)$ for each pair of values (x, y) within the range of X and Y is called the **joint probability distribution** of X and Y.

Theorem 3.7 (Joint Probability Distribution - Discrete Conditions)**.** *A bivariate function can serve as the joint probability distribution of a pair of discrete random variables X and Y if and only if its values, $f(x, y)$, satisfy the conditions:*

(1) $f(x, y) \geq 0$ *for each pair of values (x, y) within its domain*
(2) $\sum_x \sum_y f(x, y) = 1$*, where the double summation extends over all possible pairs (x, y) within its domain*

Definition 3.7 (Joint Distribution Function)**.** If X and Y are discrete random variables, the function given by

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} f(s, t) \qquad \text{for } -\infty < x < \infty$$
$$\text{for } -\infty < y < \infty$$

where $f(s, t)$ is the value of the joint probability distribution of X and Y at (s, t), is called the **joint distribution function** or the **joint cumulative distribution** of X and Y

Definition 3.8 (Joint Probability Density Function)**.** A bivariate function with values $f(x, y)$ defined over the xy-plane is called a **joint probability density function** of the continuous random variables X Y if and only if

$$P(X, Y) \in A = \iint_A f(x, y) dx dy$$

for any region A in the xy-plane

Theorem 3.8 (Joint Probability Density Function - Continuous Conditions)**.** *A bivariate function can serve as a joint probability density function of a pair of continuous random variables X and Y if its values, $f(x, y)$, satisfy the conditions:*

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

Definition 3.9 (Joint Distribution Function)**.** If X and Y are continuous random variables, the function given by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s, t) ds dt \qquad \text{for } -\infty < x < \infty$$
$$-\infty < y < \infty$$

where $f(s, t)$ is the joint probability density of X and Y at (s, t), is called the **joint distribution function of X and Y**

3.6 Marginal Distributions

Definition 3.10 (Marginal Distribution)**.** If X and Y are discrete random variables and $f(x, y)$ is the value of their joint probability distribution at (x, y), the function given by

$$g(x) = \sum_y f(x, y)$$

for each x within the range of X is called the **marginal distribution of** X. Correspondingly, the function given by

$$h(y) = \sum_x f(x, y)$$

for each y within the range of Y is called the **marginal distribution of** Y

Definition 3.11 (Marginal Density)**.** If X and Y are continuous random variables and $f(x, y)$ is the value of their joint probability density at (x, y), the function given by

$$g(x) = \int_{-\infty}^{\infty} f(x, y)dy \qquad \text{for} - \infty < x < \infty$$

is called the **marginal density of** X. Correspondingly, the function given by

$$h(x) = \int_{-\infty}^{\infty} f(x, y)dx \qquad \text{for} - \infty < y < \infty$$

is called the **marginal density of** Y.

3.7 Conditional Distributions

Definition 3.12 (Conditional Distribution)**.** If $f(x, y)$ is the value of the joint probability distribution of the discrete random variables X and Y at (x, y) and $h(y)$ is the value of the marginal distribution of Y at y, the function given by

$$f(x|y) = \frac{f(x, y)}{h(y)} \qquad h(y) \neq 0$$

for each x within the range of X is called the **conditional distribution of** X **given** $Y = y$. Correspondingly, if $g(x)$ is the value of the marginal distribution of X at x, the function given by

$$w(x|y) = \frac{f(x, y)}{g(y)} \qquad g(y) \neq 0$$

for each y within the range of Y is called the **conditional distribution of** Y **given** $X = x$.

Definition 3.13 (Conditional Density)**.** If $f(x, y)$ is the value of the joint density of the continuous random variables X and Y at (x, y) and $h(y)$ is the value of the marginal distribution of Y at y, the function given by

$$f(x|y) = \frac{f(x, y)}{h(y)} \qquad\qquad h(y) \neq 0$$

for $-\infty < x < \infty$, is called the **conditional density of** X **given** $Y = y$. Correspondingly, if $g(x)$ is the value of the marginal density of X at x, the function given by

$$w(x|y) = \frac{f(x, y)}{g(y)} \qquad\qquad g(y) \neq 0$$

for $-\infty < y < \infty$, is called the **conditional density of** Y **given** $X = x$.

Definition 3.14 (Independence of Discrete Random Variables)**.** If $f(x_1, x_2, \ldots, x_n)$ is the value of the joint probability distribution of the discrete random variables $X_1, X_2, \ldots, X_n$ at $(x_1, x_2, \ldots x_n)$ and $f_i(x_i)$ is the value of the marginal distribution of X_i at x_i for $i = 1, 2, \ldots, n$, then n random variables are **independent** if and only if

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) \cdot f_2(x_2) \cdot \ldots f_n(x_n)$$

for all $(x_1, x_2, \ldots, x_n)$ within their range

3.8 Theory in Practice

Frequency Distribution - A grouping of numerical data into classes having definite upper and lower limits
stem-and-leaf display - device for presenting quantitative data similar to histogram; generally grouped in 10s
positive skewness - long right-hand tail
negative skewness - long left-hand tail
mode - the value that appears most frequently in a data set; in a histogram it may be less general and refer to data values that are high points where the mode is a bar in a histogram that is surrounded by bars of lower frequency
bimodal - histogram exhibiting two modes
multimodal - histogram exhibiting more than two modes