

Study Notes on Probability



To have an understanding in combinations and permutations you must first have a basic understanding in probability. A probability is a numerical measure of the likelihood of the event. Probability is established on a scale from 0 to 1. A rare event has a probability close to 0; a very common event has a probability close to 1.

To solve and understand problems pertaining to probability you must know some vocabulary:

- An **experiment**, such as rolling a die or tossing a coin, is a set of trials designed to study a physical occurrence.
- An **outcome** is any result of a particular experiment. For example, the possible outcomes for flipping a coin are heads or tails.
- A **sample space** is a list of all the possible outcomes of an experiment.
- An **event** is a subset of the sample space. For example, getting heads is an event.

Counting Principles

1. Theorem (The basic principle of counting): If the set E contains n elements and the set F contains m elements, there are nm ways in which we can choose, first, an element of E and then an element of F .
2. Theorem (The generalized basic principle of counting): If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if ..., then there is a total of $n_1 \cdot n_2 \cdot \dots \cdot n_r$ possible outcomes of the r experiments.
3. Theorem: A set with n elements has 2^n subsets.
4. Tree diagrams

Permutations

1. Permutation: $n!$

$$P_r^n = \frac{n!}{(n-r)!}$$

The number of permutations of n things taken r at a time:

2. Theorem: The number of distinguishable permutations of n objects of k different types, where n_1 are alike, n_2 are alike, ..., n_k are alike and $n = n_1 + n_2 + \dots + n_k$, is

$$\frac{n!}{n_1! \times n_2! \times \dots \times n_k!}$$

Combinations

1. Combination: The number of combinations of n things taken r at a

time: $C_r^n = \frac{n!}{r!(n-r)!}$ (combinatorial coefficient; binomial coefficient)



$$(x + y)^n = \sum_{i=0}^n C_i^n x^i y^{n-i}$$

2. Binomial theorem:

3. Multinomial expansion: In the expansion of $(x_1 + x_2 + \dots + x_k)^n$, the coefficient of the

term $x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$, $n_1 + n_2 + \dots + n_k = n$, is $\frac{n!}{n_1! \times n_2! \times \dots \times n_k!}$. Therefore, $(x_1 + x_2 + \dots + x_k)^n =$

$$\sum_{n_1 + n_2 + \dots + n_k = n} \frac{n!}{n_1! \times n_2! \times \dots \times n_k!} x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$$

Note that the sum is taken over all nonnegative integers n_1, n_2, \dots, n_k such that $n_1 + n_2 + \dots + n_k = n$.

The Number of Integer Solutions of Equations

1. There are C_{r-1}^{n-1} distinct positive integer-valued vectors (x_1, x_2, \dots, x_r) satisfying $x_1 + x_2 + \dots + x_r = n$, $x_i > 0$, $i = 1, 2, \dots, r$.
2. There are C_{r-1}^{n+r-1} distinct nonnegative integer-valued vectors (x_1, x_2, \dots, x_r) satisfying $x_1 + x_2 + \dots + x_r = n$.

Axioms of Probability

Sample Space and Events

1. **Experiment:** An experiment (strictly speaking, a random experiment) is an operation which can result in two or more ways. The possible results of the experiment are known as outcomes of the experiment. These outcomes are known in advance. We cannot predict which of these possible outcomes will appear when the experiment is done.
2. **Sample Space:** The set containing all the possible outcomes of an experiment as its element is known as sample space. It is usually denoted by S.
3. **Event:** An event is a subset of the sample space S.

Illustration 1:

Let us consider the experiment of tossing a coin. This experiment has two possible outcomes: heads (H) or tails (T)

- sample space (S) = {H, T}

We can define one or more events based on this experiment. Let us define the two events A and B as:

A: heads appears



B: tails appear

It is easily seen that set A (corresponding to event A) contains outcomes that are **favourable to event A** and set B contains outcomes favourable to event B. Recalling that $n(A)$ represents the number of elements in set A, we can observe that

$n(A)$ = number of outcomes favourable to event A

$n(B)$ = number of outcomes favourable to event B

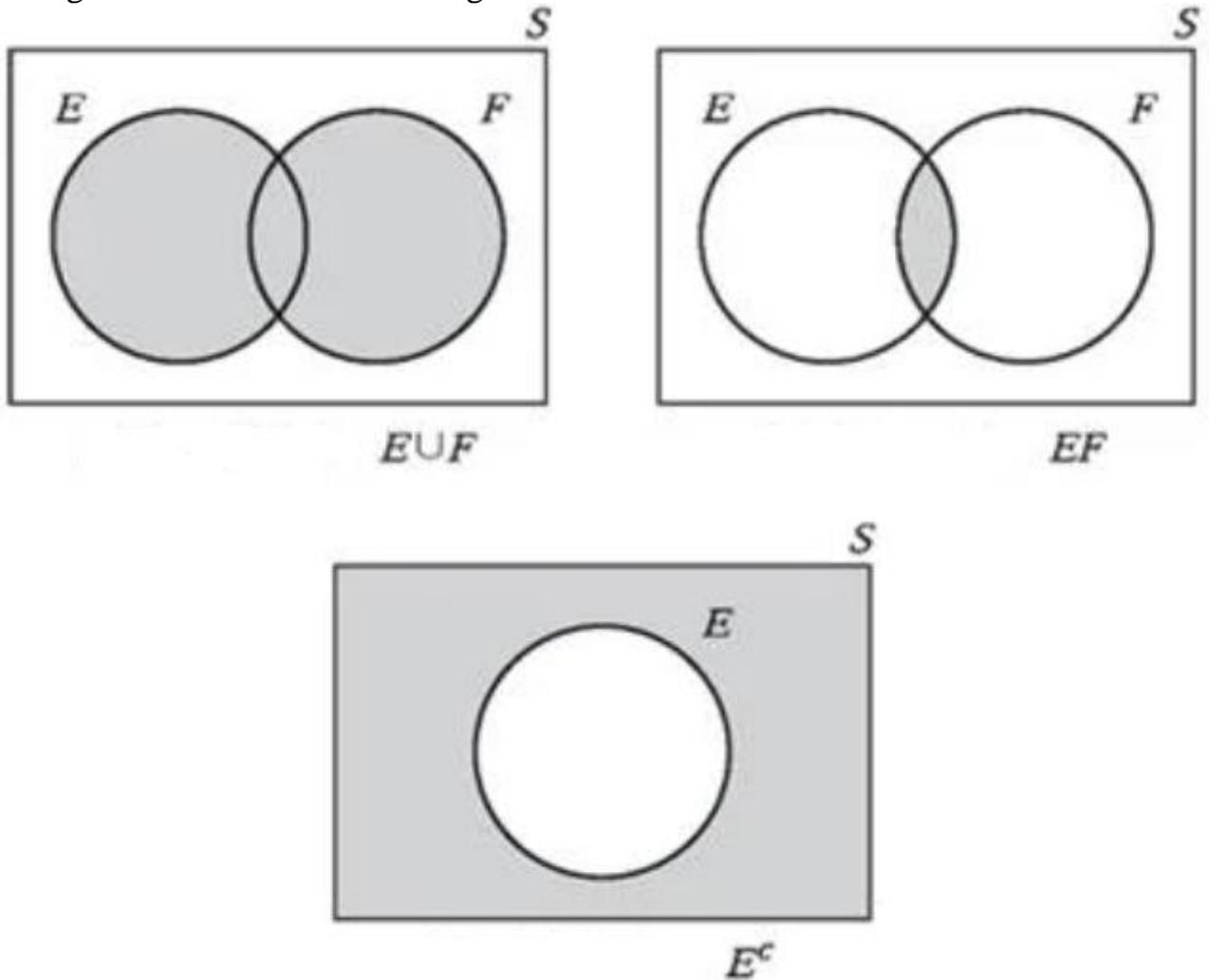
$n(S)$ = number of possible outcomes

Here, in this example, $n(A) = 1$, $n(B) = 1$, and $n(S) = 2$.

- Set theory concepts: set, element, roster method, rule method, subset, null set (empty set).
- Complement: The complement of an event A with respect to S is the subset of all elements of S that are not in A. We denote the complement of A by the symbol A' (A^c).
- Intersection: The intersection of two events A and B, denoted by the symbol $A \cap B$, is the event containing all elements that are common to A and B.
Two events A and B are mutually exclusive, or disjoint, if $A \cap B = \varnothing$ that is, if A and B have no elements in common.



- The union of the two events A and B , denoted by the symbol $A \cup B$, is the event containing all the elements that belong to A or B or both.



- Venn diagram:
 - Sample space of an experiment: All possible outcomes (points)
 - Events: subsets of the sample space
- Impossible events (impossibility): Φ ; sure events (certainty): S .

Set Algebra	Probability
Set Algebra	Event
Universal Set	Sample space
Element	Outcome



DeMorgan's laws:

$$\left(\bigcup_{i=1}^n E_i \right)^c = \bigcap_{i=1}^n E_i^c$$

$$\left(\bigcap_{i=1}^n E_i \right)^c = \bigcup_{i=1}^n E_i^c$$

Axioms of Probability

1. Probability axioms:

(1) $0 \leq P(A) \leq 1$;

(2) $P(S) = 1$;

(3) $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$ If $\{A_1, A_2, \dots\}$ is a sequence of mutually exclusive events.

2. Equally likely outcomes: the probabilities of the single-element events are all equal.

A few sample events are said to be equally likely if there is no reason for one event to occur in preference to any other event.

Basic Theorems

(1) $0 \leq P(A) \leq 1$;

(2)
$$P(A) = \frac{N(A)}{N(S)}$$
;

(3) Complementary events: $P(\bar{A}) = 1 - P(A)$;

(4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$: inclusion-exclusion principle

(5) If A_1, A_2, A_n is a partition of sample space S , then

$$\begin{aligned} &P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) \\ &= P(A_1) + P(A_2) + \dots + P(A_n) \\ &= P(S) \\ &= 1 \end{aligned}$$

(6) if a and a' are complementary events, then $P(A) + P(A') = 1$.

Conditional Probability and Independence

Conditional Probability

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

1. Conditional probability:



Consider two events A and B defined on a sample space S. The probability of occurrence of event A given that event B has already occurred is known as the conditional probability of A relative to B.

$P(A/B) = P(A \text{ given } B) = \text{probability of occurrence of } A \text{ assuming that } B \text{ has occurred}$

It implies that the outcomes favourable to B become the possible outcomes and hence outcomes favourable to $P(A/B)$ are outcomes common to A and B.

Let $n(B)=a$ and $n(A \cap B) = b$. Suppose that the event B occurs. Then there are exactly a sample points and these points can be the sample space for the other event A. The event A in this sample space consists of b sample points common to A and B.

Therefore, the probability of A in this sample space = b/a .

Thus, the probability of A under the assumption that event B has occurred is:

$$P(A / B) = \frac{b}{a} = \frac{n(A \cap B)}{n(B)} = \frac{n(A \cap B) / n(S)}{n(B) / n(S)}$$

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

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

and similarly,

The above result is known as **conditional probability theorem**.

2. If in an experiment the events A and B can both occur, then $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$.

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

The multiplication rule: $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$.

3. Partition: Let $\{B_1, B_2, \dots, B_n\}$ be a set of nonempty subsets of the sample space S of an experiment. If the events B_1, B_2, \dots, B_n are mutually exclusive and $B_1 \cup B_2 \cup \dots \cup B_n = S$, the set $\{B_1, B_2, \dots, B_n\}$ is called a partition of S.

4. Theorem of total probability: If B_1, B_2, \dots is a partition of S, and A is any event,

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)$$

then



Total Probability Theorem: The probability that one of several mutually exclusive events A_1, A_2, \dots, A_n will happen is the sum of the probabilities of the separate events. i.e.

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

5. Bayes' Theorem: If B_1, B_2, \dots is a partition of S , and A is any event, then

$$P(B_i | A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^{\infty} P(B_i)P(A|B_i)}$$

Independence

1. Independent events: If A, B are independent events $\leftrightarrow P(A \cap B) = P(A)P(B)$.
2. Theorem: If A and B are independent, then A and \bar{B} ; \bar{A} and B are independent.
3. The events A, B , and C are called independent if $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$, $P(A \cap B \cap C) = P(A)P(B)P(C)$. If A, B , and C are independent events, we say that $\{A, B, C\}$ is an independent set of events.
4. The set of events $\{A_1, A_2, \dots, A_n\}$ is called independent if for every

subset $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}, k \geq 2$ of $\{A_1, A_2, \dots, A_n\}$,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k})$$

Important Results:

Following results are easily derived from the definition of independent events.

1. A and B are independent if $P(B|A) = P(B)$ and $P(A|B) = P(A)$
2. If A and B are independent, then $P(A \cap B) = P(A)P(B)$
3. A set of events A_1, A_2, \dots, A_n are said to be pair-wise independent if $P(A_i \cap A_j) = P(A_i)P(A_j) \forall i \neq j$.
4. The events A_1, A_2, \dots, A_n are mutually independent (or simply independent) if and only if the multiplication rule:

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1)P(A_2) \dots P(A_k) \dots \quad (1)$$
 holds for every t triples of events, $k = 2, 3, \dots, n$. If (1) holds for $k = 2$ and may not hold



for $k = 3, 4, \dots, n$ then events A_1, A_2, \dots, A_n are said to be pair wise independent. Thus mutually independent events are pair wise independent but converse is not true.

5. If A and B are mutually independent events such that $P(A) \neq 0$ and $P(B) \neq 0$, then the events A and B have at least one common sample point (i.e. they cannot be mutually exclusive). Or in general, **mutually exclusive events are dependent events.**

Distribution Functions and Discrete Random Variables

Distribution Functions

1. Cumulative Distribution Functions (*cdf*): $F_X(t) = P(X \leq t)$, $-\infty < t < \infty$.

2. $F_X(t)$ is non-decreasing; $0 \leq F_X(t) \leq 1$; $\lim_{t \rightarrow -\infty} F_X(t) = 0$, $\lim_{t \rightarrow \infty} F_X(t) = 1$

If $c < d$, then $F_X(c) \leq F_X(d)$; $P(c < X \leq d) = F_X(d) - F_X(c)$; $P(X > c) = 1 - F_X(c)$.

The *cdf* of a discrete random variable: a step function.

Special Discrete Distributions

Bernoulli and Binomial Random Variables

1. Bernoulli trials: an experiment with two different possible outcomes

Bernoulli random variable X with parameter p , p is the probability of a success.

$$p_X(x) = p^x(1-p)^{1-x} = p^x q^{1-x}, \text{ for } x \in R_X = \{0, 1\}$$

Expected value: $E[X] = \mu_x = p$; variance: $Var[X] = \sigma_x^2 = p(1-p) = pq$

2. Binomial distribution: number of successes to occur in n repeated, independent Bernoulli trials

Binomial random variable Y with parameters n and p

$$p_Y(y) = C_y^n p^y q^{n-y}, \text{ for } y \in R_Y = \{0, 1, 2, \dots, n\}$$

Expected value: $E(Y) = \mu_Y = np$; Variance: $Var[Y] = \sigma_Y^2 = npq$

Multinomial Random Variables

1. Multinomial trials: an experiment with $k \geq 2$ different possible outcomes

2. Multinomial distribution: n independent multinomial trials

Multinomial random variable X_1, X_2, \dots, X_k with parameters n, p_1, p_2, \dots, p_k ; X_i : the number of i th

outcomes; $\sum_{i=1}^k x_i = n$; $\sum_{i=1}^k p_i = 1$; $R_{X_i} = \{0, 1, \dots, n\}$



$$P_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) \\ = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}, \text{ for } (x_1, x_2, \dots, x_k) \in R_{X_1, X_2, \dots, X_k}$$

Geometric distribution: trial number of the first success to occur in a sequence of independent Bernoulli trials

Geometric random variable N with parameter p

$$p_N(n) = pq^{n-1}, \text{ for } n \in R_N = \{1, 2, 3, \dots\}$$

Geometric series:

$$1 + q + q^2 + \dots = \sum_{i=0}^{\infty} q^i = \frac{1}{1-q}, \text{ if } |q| < 1$$

Expected value: $E(N) = \mu_N = \frac{1}{p}$; variance: $Var[N] = \sigma_N^2 = \frac{q}{p^2}$

Negative binomial distribution: trial number of the r th success to occur

Negative binomial random variable N , with parameters r, p

$$P_{N_2}(n) = p C_1^{n-1} p q^{n-2} \\ = (n-1) p^2 q^{n-2}, \text{ for } n \in R_{N_2} = \{2, 3, 4, \dots\}$$

$$P_{N_r}(n) = C_{r-1}^{n-1} p^r q^{n-r}, \text{ for } n \in R_{N_r} = \{r, r+1, r+2, \dots\}$$

Expected value: $E(N_r) = \mu_{N_r} = \frac{r}{p}$; variance: $Var[N_r] = \sigma_{N_r}^2 = \frac{rq}{p^2}$



Poisson Distribution

1. The Poisson probability function:

$$P_K(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad \text{for } k = 0, 1, 2, 3, \dots$$

Poisson random variable K with parameter λ

Expected value: $E(K) = \mu_K = \lambda$; variance: $Var[K] = \sigma_K^2 = \lambda$

2. The Poisson approximation to the binomial: If X is a binomial random variable with parameters n and $p = \lambda/n$, then

$$\lim_{n \rightarrow \infty} P_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad \text{for } x \in R_X = \{0, 1, 2, \dots\}$$

Continuous Random Variables

Probability Density Functions

1. Densities

2. Probability density function (*pdf*) for a continuous random variable X : $f_x(x)$

$$f_x(x) \geq 0$$

For all

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

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx;$$

$$P(X \leq b) = \int_{-\infty}^b f_X(x) dx = F_X(b);$$

$$P(X \geq a) = \int_a^{\infty} f_X(x) dx;$$

$$f_X(\infty) = f_X(-\infty) = 0;$$

$$P(X = x) = 0;$$

$$P(X \approx x) = f_X(x) dx$$



Cumulative Distribution Functions (cdf)

1.

$$F_X(t) = P(X \leq t)$$
$$= \begin{cases} \sum_{x \leq t} P_X(x), & \text{for a discrete random variable } X \\ \int_{-\infty}^t f_X(x) dx, & \text{for a continuous random variable } X \end{cases}$$

, $-\infty < t < \infty$

2. The *cdf* of a discrete random variable: a step function; the *cdf* of a continuous random variable: a continuous function

3. The probability function of a discrete random variable: size of the jump in $F_X(t)$; the *pdf* of a

continuous random variable: $f_X(t) = \frac{d}{dt} F_X(t)$

Expectations and Variances

1. Definition: If X is a continuous random variable with *pdf* $f_X(x)$, the expected value of X is

defined by $E[X] = \mu_X = \int_{-\infty}^{\infty} xf_X(x) dx$.

Example: In a group of adult males, the difference between the uric acid value and 6, the

standard value, is a random variable X with the following *pdf*: $f_X(x) = \frac{27}{490} (3x^2 - 2x)$ if $2/3 < x < 3$. Calculate the mean of these differences for the group.

2. Theorem: Let X be a continuous random variable with *pdf* $f_X(x)$; then for any

function h : $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx$.

3. Corollary: Let X be a continuous random variable with *pdf* $f_X(x)$. Let h_1, h_2, \dots, h_n be real-valued functions and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers. Then $E[\alpha_1 h_1(X) + \alpha_2 h_2(X) + \dots + \alpha_n h_n(X)] = \alpha_1 E[h_1(X)] + \alpha_2 E[h_2(X)] + \dots + \alpha_n E[h_n(X)]$.

4. Definition: If X is a continuous random variable with $E[X] = \mu_x$, then $Var[X]$ and σ_x , called the variance and standard deviation of X , respectively, are defined by



$$\text{Var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$$

$$\sigma_X = \sqrt{\text{Var}[X]}$$

Special Continuous Distributions

Uniform Random Variable

1. Density of a uniformly distributed random variable: $f_X(x) = \frac{1}{b-a}$, for $a \leq x \leq b$
2. The *cdf* of a uniformly distributed random variable:

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x \geq b \end{cases}$$

3. The expected value and variance of the uniform random variable: $\mu_X = E[X] = \frac{a+b}{2}$

$$\text{Var}[X] = \sigma_X^2 = \frac{(b-a)^2}{12}$$

The Exponential Distribution

$$F_{T_1}(t) = 1 - e^{-\lambda t},$$

1. The exponential probability law with parameter λ : $f_{T_1}(t) = \lambda e^{-\lambda t}$, for $t \geq 0$.

T_1 is the time of occurrence of the first event in a Poisson process with parameter λ , starting at an arbitrary origin $t=0$; $\{X_{\lambda t} = 0\} \equiv \{T_1 > t\}$, Poisson process $X_{\lambda t}$ is the number of events to occur with parameter $\mu = \lambda t$.

2. The expected value and variance of the exponential random variable:

$$\mu_{T_1} = E[T_1] = \frac{1}{\lambda}; \text{Var}[T_1] = \sigma_{T_1}^2 = \frac{1}{\lambda^2}$$



The Erlang Distribution

$$F_{T_2}(t) = 1 - (1 + \lambda t)e^{-\lambda t},$$

1. The *cdf* and *pdf* for T_2 : $f_{T_2}(t) = \lambda^2 t e^{-\lambda t}$, for $t > 0$.

T_2 is the time of occurrence of the second event in a Poisson process with parameter λ , starting at an arbitrary origin $t=0$; $\{X_{\lambda t} \leq 1\} \equiv \{T_2 > t\}$, Poisson process $X_{\lambda t}$ is the number of events to occur with parameter $\mu = \lambda t$?

$$F_{T_r}(t) = \sum_{k=r}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

$$f_{T_r}(t) = \frac{\lambda^r t^{r-1}}{(r-1)!} e^{-\lambda t}, \text{ for } t > 0.$$

2. The Erlang probability law with parameters r and λ :

T_r is the time of occurrence of the r th event in a Poisson process with parameter λ , starting at an arbitrary origin $t=0$; $\{X_{\lambda t} \leq r - 1\} \equiv \{T_r > t\}$, $X_{\lambda t}$ is the number of events to occur.

3. The expected value and variance of the Erlang random variable:

$$\mu_{T_r} = E[T_r] = \frac{r}{\lambda}; \text{Var}[T_r] = \sigma_{T_r}^2 = \frac{r}{\lambda^2}$$

The Gamma Distribution

$$f_U(u) = \frac{\lambda^n u^{n-1}}{\Gamma(n)} e^{-\lambda u}, \text{ for } u > 0, n > 0, \lambda > 0.$$

1. The gamma probability law with parameters n and λ :

Gamma function: $\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx, n > 0$; $\Gamma(n) = (n-1) \Gamma(n-1)$; $\Gamma(n) = (n-1)!$, if n is a positive integer.

The Erlang random variable is a particular case of a gamma random variable, where n is restricted to the integer values $r=1, 2, 3, \dots$



2. The expected value and variance of the gamma random variable:

$$E[U] = \mu_U = \frac{n}{\lambda}; \text{Var}[U] = \sigma_U^2 = \frac{n}{\lambda^2}$$

The Normal (Gaussian) Distribution

1. The normal probability law with parameters μ and σ : $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$, for $-\infty < x < \infty, \sigma > 0$.

2. The expected value and variance of the normal random variable:

$$\mu_X = E[X] = \mu; \text{Var}[X] = \sigma_X^2 = \sigma^2$$

3. Standard normal random variable (the unit normal): $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, $\mu=0, \sigma=1$.

4. The new random variable: If X is normal with mean μ and variance σ^2 , then $X+b$ is normal with mean $\mu+b$ and variance σ^2 ; aX is normal with mean $a\mu$ and variance $a^2\sigma^2$; $aX+b$ is normal with mean $a\mu+b$ and variance $a^2\sigma^2$.

5. Switching from a non-unit normal to the unit normal: $Z = \frac{X - \mu}{\sigma}$, $\mu_z=0, \sigma_z=1$.

6. The probability that a normal random variable is with k standard deviations of the mean: $P(\mu - k\sigma \leq X \leq \mu + k\sigma) = P(-k \leq X^* \leq k)$, X^* is the unit normal.

7. The normal approximation to the binomial: A binomial distribution X with parameters n, p where n is large, then X is approximately normal with $\mu = np, \sigma^2 = npq$.

Jointly Distributed Random Variables

Joint Densities

1. Jointly distributed random variable: If the observed values for two or more random variables are simultaneously determined by the same random mechanism.



2. The discrete random variables X_1, X_2 : the probability function

is $p_{X_1, X_2}(x_1, x_2); p_{X_1, X_2}(x_1, x_2) \geq 0$ for all x_1, x_2 ; $\sum_R p_{X_1, X_2}(x_1, x_2) = 1$. The continuous random variables X_1, X_2 : the pdf is positive; $f_{X_1, X_2}(x_1, x_2) \geq 0$ for all x_1, x_2 ; $\iint_R f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = 1$;

the probability is given by integrating $f_{X_1, X_2}(x_1, x_2)$.

3. The joint density of independent random variables: The random variables X, Y are independent if and only if $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, when X, Y are discrete; $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, when X, Y are continuous.

4. Uniform joint densities: If X and Y are jointly uniform on a region, the

joint pdf is $f_{X,Y}(x,y) = \frac{1}{\text{area of region}}$, for (x,y) in the region.

The joint density of independent uniform random variables: $P(\text{event}) = \frac{\text{favorable area}}{\text{total area}}$

Marginal Densities

1. The discrete random variables X_1, X_2 with the probability function $p_{X_1, X_2}(x_1, x_2)$ and range R_{X_1, X_2} , the marginal probability function

for X_1 is $p_{X_1}(x_1) = \sum_{x_2} p_{X_1, X_2}(x_1, x_2)$ for $x_1 \in R_{X_1}$. R_{X_1} is the marginal range for X_1 , the set of first elements of $(x_1, x_2) \in R_{X_1, X_2}$.

2. The continuous random variables X_1, X_2 with pdf $f_{X_1, X_2}(x_1, x_2)$ and range R_{X_1, X_2} , the

marginal pdf for X_1 is $f_{X_1}(x_1) = \int_{x_2} f_{X_1, X_2}(x_1, x_2) dx_2$ for $x_1 \in R_{X_1}$. R_{X_1} is the marginal range

for X_1 , the set of first elements of $(x_1, x_2) \in R_{X_1, X_2}$.

Sums of Independent Random Variables

1. The sum of independent binomials with a common p : X_1 with parameters n_1, p , X_2 with parameters n_2, p , then $X_1 + X_2$ with parameters $n_1 + n_2, p$

2. The sum of independent Poissons: X_1 with parameter λ_1 , X_2 with parameters λ_2 , then $X_1 + X_2$ with parameters $\lambda_1 + \lambda_2$

3. The sum of independent exponentials with a common λ : X_1, \dots, X_n with parameter λ , then $X_1 + \dots + X_n$ has a gamma (Erlang) distribution with parameters n, λ .



4. The density of the sum of two arbitrary independent random variables: the convolution of the individual *pdf*,

$$f_{X+Y}(z) = \int_{\mathbb{R}_x} f_X(x) f_Y(z-x) dx = \int_{\mathbb{R}_y} f_X(z-y) f_Y(y) dy$$

5. The sum of independent normal: X_1 with mean μ_1 and variance σ_1^2 , X_2 with mean μ_2 and variance σ_2^2 , then X_1+X_2 with mean $\mu_1+\mu_2$ and variance $\sigma_1^2 + \sigma_2^2$



Prepare for BBA & IPM Entrance Exams



Live Classes by
Top Faculty



Comprehensive Study
Material



Daily Study Plan



Latest Pattern Test Series



Complete Doubt
Resolution



Regular Assessments with
Report Card

