STAT-2920 Intro to Probability Review

Aidan Richer

November 2025

This paper contains a cumulative review of all definitions and theorems within the lecture notes, excluding any examples.

Contents

1 Chapter 1: Combinatorial Analysis

1.1 The Basic Principle of Counting

Theorem. *Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of* m *possible outcomes and if, for each outcome of experiment 1, there are* n *possible outcomes of experiment 2, then together there are* $m \times n$ *possible outcomes of the two experiments.*

1.2 Permutations

Frequently, we are interested in situations where the outcomes are the different ways in which a group of objects can be ordered or arranged. Different arrangements like these are called permutations.

Definition 1. A permutation is a distince arrangement of n different elements of a set.

Theorem. *The number of permutations of* n *distinct objects taken* r *at a time is*

$$\mathrm{P}_r^n = \frac{n!}{(n-r)!}$$

for $r = 1, 2, \ldots, n$

Theorem. *The number of permutations of* n *objects of which* n_1 *are of one kind,* n_2 *are of a second kind,* $\ldots, n_k$ *are of a* k^{th} *kind, and* $n_1 + n_2 + \cdots + n_k = n$ *is*

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

1.3 Combinations

Definition 2. A combination is a selection og r objects taken from n distincy objects without regard to the order of selection.

Theorem. *The number of combinations of n distinct objects taken r at a time is*

$$\mathrm{C}_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

for $r = 0, 1, 2, \ldots, n$

Theorem. *The number of ways in which a set of n distinct objects can be partitioned into k subsets with n_1 objects in the first subset, n_2 objects in the second subset, …, and n_k objects in the k^{th} subset is*

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! \times n_2! \times \cdots \times n_k!}, \quad n = n_1 + n_2 + \cdots + n_k$$

1.4 Binomial Coefficients

Definition 3. The coefficient of $x^{n-r}y^r$ in the binomial expansion of $(x+y)^n$ is called the binomial coefficient $\binom{n}{r}$.

Theorem.

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r$$

for any positive integer n.

The calculation of binomial coefficients can often be simplified by making use of the three theorems that follow.

Theorem. *For any positive integers n and $r = 0, 1, 2, \ldots, n$,*

$$\binom{n}{r} = \binom{n}{n-r}$$

Theorem. *For any positive integer n and $r = 1, 2, \ldots, n-1$, we have*

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

Theorem.

$$\sum_{r=0}^{k} \binom{m}{r} \binom{n}{k-r} = \binom{m+n}{k}$$

2 Chapter 2: Probability

2.1 Review of Set Theory

Definition 4. A set is an unordered collection of objects called elements of the set.

Definition 5. Two sets are equal if they contain exactly the same elements. We say

$$a \in \mathcal{A} \text{ means that element } a \text{ belongs to the set } \mathcal{A}$$

$$a \notin \mathcal{A} \text{ means the element } a \text{ does not belong to the set } \mathcal{A}$$

$$|\mathcal{A}| \text{ means the size, or the cardinality of a set } \mathcal{A}$$

Definition 6. A set A is a subset of a set B if every element of A is also an element of B. We write

$$A \subseteq B$$

If A is not a subset of B, we write

$$A \nsubseteq B$$

Definition 7. $S = \Omega$ is the universal set i.e. the collection of all points of interest for the present situation. The empty set, denoted $\emptyset$ is the set with no elements.

Law (Commutative Laws).

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

Law (Associative Laws).

$$(A \cup B) \cup C = A \cup (B \cup C)$$

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

Law (Distributive Laws).

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

Law (Basic Sample Space Laws).

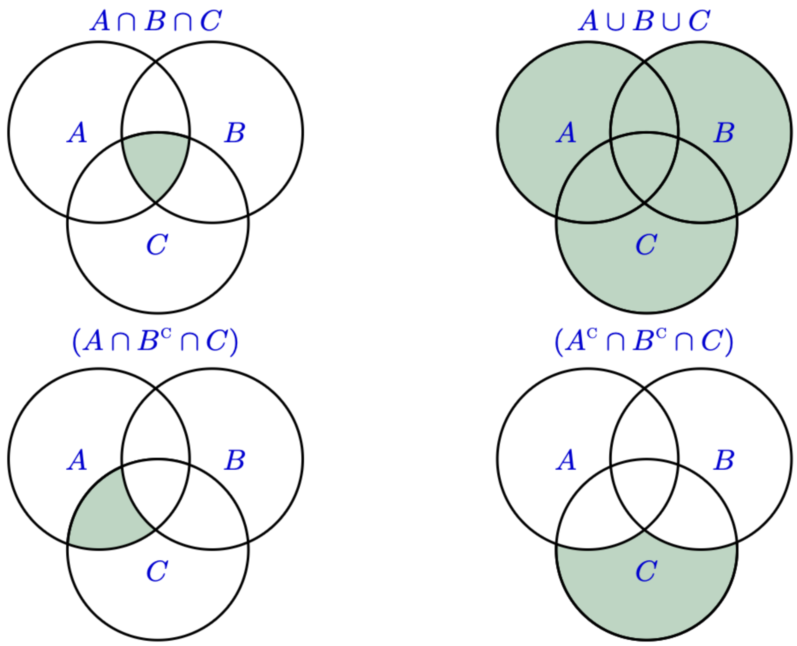
$$A \cap A^C = \emptyset$$

$$A \cup A^C = S$$

Law (De Morgan's Laws).

$$(A \cup B)^C = A^C \cap B^C$$

$$(A \cap B)^C = A^C \cup B^C$$



2.2 The Probability of an Event

Theorem. *If A is an event in a discrete sample space S, then $\mathbb{P}(A)$ equals the sum of the probabilities of the individual outcomes comprising A.*

Theorem. *If an experiment can result in any one of N different equally likely outcomes, and if n of these outcomes together constitute event A, then the probability of event A is*

$$\mathbb{P}(A) = \frac{n}{N}$$

2.3 Properties of Probability in Sets

Theorem (Properties)**.** *A probability $\mathbb{P}$ satisfies the following properties:*

1. *For each event A,* $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$
2. *The probability of the null set is zero; that is,* $\mathbb{P}(\emptyset) = 0$
3. $\mathbb{P}(A) \leq \mathbb{P}(B)$ whenever $A \subset B$
4. *For each A,* $0 \leq \mathbb{P}(A) \leq 1$

Theorem. *For any events A and B, we have*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Theorem. *If A, B and C are any three events in a sample space S, then*

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

2.4 Conditional Probability

Definition 8. If A and B are any two events in a sample space S and $\mathbb{P}(A) > 0$, the conditional probability of B given A is

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

Theorem. *If A and B are any two events in a sample space S and $\mathbb{P}(A) \neq 0$, then*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B \mid A)$$

Theorem. *If $A, B,$ and C are any three event in a sample space S such that $\mathbb{P}(A \cap B) \neq 0$, then*

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B \mid A) \cdot \mathbb{P}(C \mid A \cap B)$$

2.5 Independent Events

Definition 9. Two events A and B are independent if and only if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Theorem. *If A and B are independent, then A and B^C are also independent.*

Definition 10. Events $A_1, A_2, \ldots, A_k$ are independent if and only if the probability of the intersections of any $2, 3, \ldots,$ or k of these events equals the product of their respective probabilities. For three events, $A, B,$ and C, for example, independence requires that

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

$$\mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C)$$
$$\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C)$$
$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$$

2.6 Law of Total Probability

Theorem. *If the events $B_1, B_2, \ldots,$ and B_k constitute a partition of the sample space S and $\mathbb{P}(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S*

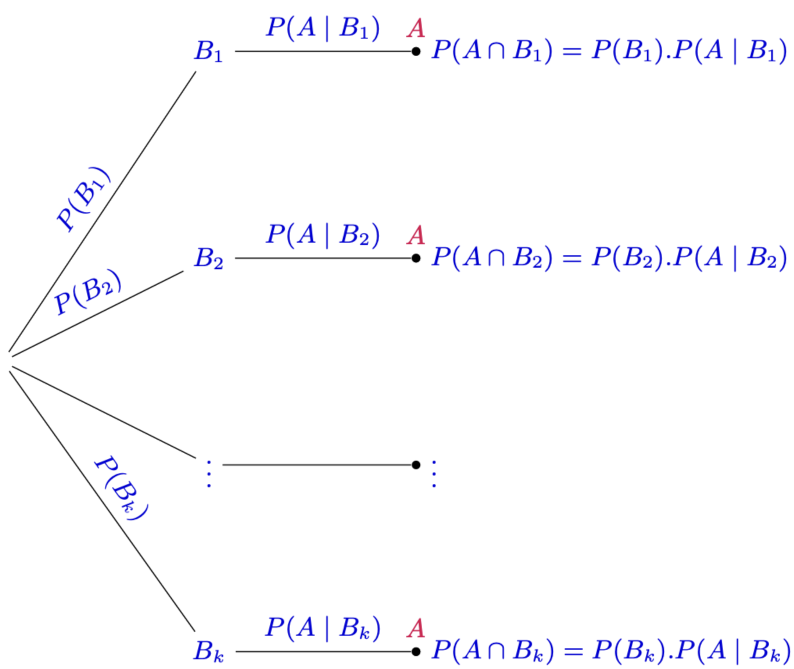
$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{k} \mathbb{P}(B_i) \cdot \mathbb{P}(A \mid B_i)$$

2.7 Bayes' Theorem

Theorem. *If the events $B_1, B_2, \ldots,$ and B_k constitute a partition of the sample space S and $\mathbb{P}(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S, such that $\mathbb{P}(A) \neq 0$*

$$\mathbb{P}(B_r \mid A) = \frac{\mathbb{P}(B_r) \cdot \mathbb{P}(A \mid B_r)}{\sum_{i=1}^{k} \mathbb{P}(B_i) \cdot \mathbb{P}(A \mid B_i)}$$

for $r = 1, 2, \ldots, k$



3 Chapter 3: Probability Distribution and Probability Density

3.1 Random Variables

In this chapter, random variables are denoted by capital letters and their values by the corresponding lowercase letter. For instance, we shall write x to denote a value of the random variable X.

Definition 11. 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.

Definition 12. X is a discrete random variable if its range is finite or countable infinite.

3.2 Probability Mass Function

Definition 13 (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. $f(x)$ is also called the probability mass function (pmf).

Theorem. *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.*

3.3 Cumulative Distribution Function

Definition 14. If X is a discrete random variable, the function given by

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

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

Theorem. *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.

Theorem. *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*

$$f(x_i) = F(x_i) - F(x_{i-1}) \qquad \textit{for } i = 2, 3, 4, \ldots, n$$

3.4 Probability Density Functions

Definition 15. A function with values $f(x)$, defined over the set of all real numbers, is called a probability density function (pdf) of the continuous random variable X if and only if

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

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

Theorem. *If X is a continuous random variable and a and b are real constants with $a \leq b$, then*

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b)$$

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

3.5 Distribution Functions

Definition 16. If X is a continuous random variable and the value of its probability density at t is $f(t)$, then the function given by

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

Theorem. *If $f(x)$ and $F(x)$ are the values of the probability density and the distribution function of X at x, then*

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

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

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

where the derivative exists.

4 Chapter 4: Mathematical Expectation

4.1 The Expected Value of a Random Variable

Definition 17. Let X be a random variable.

- If X is a discrete and $f(x)$ is the value of its probability distribution at x, the expected value of X is
$$\mathbb{E}[X] = \sum_x x \cdot f(x)$$
- If X is a continuous and $f(x)$ is the value of its probability density at x, the expected value of X is
$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x \cdot f(x)dx$$

Theorem.
- *If X is a discrete random variable and $f(x)$ is the value of its probability distribution at x, the expected value of $g(X)$ is given by*
$$\mathbb{E}[g(X)] = \sum_x g(x) \cdot f(x)$$
- *Correspondingly, if X is a continuous random variable and $f(x)$ is the value of its probability density at x, the expected value of $g(X)$ is given by*
$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) \cdot f(x)dx$$

Definition 18. If X and Y are two random variables and a and b are any real numbers, then

- $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$
- $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$

Theorem. *If $c_1, c_2, \ldots,$ and c_n are constants, then*
$$\mathbb{E}\left[\sum_{i=1}^{n} c_i g_i(X)\right] = \sum_{i=1}^{n} c_i \mathbb{E}[g_i(X)]$$

4.2 Moments

Definition 19 (Moment)**.** The r^{th} moment about the origin of a random variable X, denoted by μ'_r, is the expected value of X^r; symbolically
$$\mu'_r = \mathbb{E}(X^r) = \sum_x x^r \cdot f(x)$$
for $r = 0, 1, 2, 3, \ldots,$ when X is discrete, and
$$\mu'_r = \mathbb{E}(X^r) = \int_{-\infty}^{\infty} x^r \cdot f(x)dx$$
when X is continuous.

Definition 20 (Mean)**.** μ_1', is called the <u>mean</u> of the distribution of X, or simply the mean of X, and it is denoted simply by μ.

Definition 21 (Moments about the mean)**.** The r^{th} moment about the mean of a random variable X, denoted by μ_r, is the expected value of $(X-\mu)^r$; symbolically

$$\mu_r = \mathbb{E}[(X-\mu)^r] = \sum_x (x-\mu)^r \cdot f(x)$$

for $r = 0, 1, 2, 3, \ldots$, when X is discrete, and

$$\mu_r = \mathbb{E}[(X-\mu)^r] = \int_{-\infty}^{\infty} (x-\mu)^r \cdot f(x)dx$$

when X is continuous.

Definition 22 (Variance)**.** μ_2, is called the <u>variance</u> of the distribution of X, or simply the variance of X, and it is denoted simply by $\sigma^2, \sigma_X^2, Var(X)$, or $V(X)$. The positive square root of the variance, σ, is called the <u>standard deviation</u> of X. Mathematically, we express σ^2 as

$$\sigma^2 = \mu_2' - (\mu_1')^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Theorem. *If X has variance σ^2, then*

$$Var(aX+b) = a^2 Var(X) = a^2\sigma^2$$

4.3 Moment Generating Function

Definition 23. The moment generating function of a random variable X, where it exists, is given by

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \sum_x e^{tx} \cdot f(x)$$

when X is discrete, and

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)dx$$

when X is continuous.

Theorem.

$$\frac{d^r}{dt^r} M_X(t) = \mu_r'$$

Theorem. *If a and b are constants, then*

1. $$M_{X+a}(t) = e^{at} \cdot M_X(t)$$

2. $$M_{bX}(t) = M_X(bt)$$

3. $$M_{\frac{X+a}{b}}(t) = \mathbb{E}\left[e^{(\frac{X+a}{b})t}\right] = e^{\frac{a}{b}t} \cdot M_X\left(\frac{t}{b}\right)$$

5 Chapter 5: Special Probability Distributions

5.1 Discrete Uniform Distribution

Definition 24. A random variable X has a discrete uniform distribution and it is referred to as a discrete uniform random variable if and only if its probability distribution is given by

$$f(x) = \frac{1}{k} \qquad \text{for } x = x_1, x_2, \ldots, x_k$$

where $x_i \neq x_j$ when $i \neq j$

5.2 Bernoulli Distribution

Definition 25. A random variable X has a Bernoulli distribution and it is referred to as a Bernoulli random variable if and only if its probability distribution is given by

$$f(x;\theta) = \theta^x(1-\theta)^{1-x} \qquad \text{for } x = 0, 1$$

Thus,

$$\mathbb{P}(X = 0) = 1 - \theta \text{ and } \mathbb{P}(X = 1) = \theta$$

for some $0 \leq \theta \leq 1$. We will often simply write $X \sim \text{Ber}(\theta)$ to indicate such a random variable.

Theorem (Properties of a Bernoulli Distribution). *For a Bernoulli random variable $X \sim Ber(\theta)$*

- *The cdf is given by*

$$F_X(x;\theta) = \begin{cases} 0 & x < 0 \\ 1-\theta & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

- *The moment generating function is*

$$M_X(t) = (1-\theta) + \theta e^t$$

- *Expectation is*

$$\mathbb{E}[X] = \theta$$

- *Variance is*

$$\mathbb{V}ar(X) = \theta(1-\theta)$$

5.3 Binomial Distribution

Definition 26. Consider n independent trials, each of which has a probability of success θ, and probability of failure $1-\theta$. If X represents the number of successes that occur in the n trials, then X is said to be a binomial random variable with parameters (n, θ). It is denoted as

$$X \sim \text{Bin}(n, \theta)$$

Theorem (Properties of a Binomial Distribution)**.** *For a binomial random variable* $X \sim \text{Bin}(n, \theta)$

- *The probability mass function is*

$$f(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

- *Moment generating function (mgf) is*

$$M_X(t) = (1 - \theta + \theta e^t)^n$$

- *Expectation is*

$$\mathbb{E}[X] = n\theta$$

- *Variance is*

$$\mathbb{V}ar(X) = n\theta(1-\theta)$$

5.4 Negative Binominal Distribution

In connection with repeated Bernoulli trials, we are sometimes interested in the number of the trial on which the k^{th} success occurs.

Definition 27. Suppose that independent trials, each having probability of success $0 < \theta < 1$, are performed until a total of r successes are accumulated. If X is the number of trials required, then X is said to be a negative binomial random variable. It is denoted as

$$X \sim \text{NB}(r, \theta)$$

The last trial must necessarily result in a success, and there must be $r-1$ more success in the first $x-1$ trials.

Theorem (Properties of a Negative Binomial Distribution)**.** *For a negative binomial random variable*
$X \sim \text{NB}(r, \theta)$

- *The probability distribution of X is*

$$f(x; r, \theta) = \mathbb{P}(X = x) = \binom{x-1}{r-1} \theta^r (1-\theta)^{x-r}$$

- *Moment generating function (mgf) is*

$$M_X(t) = \left(\frac{\theta e^t}{1-(1-\theta)e^t}\right)^r \qquad \text{for } t < -\log(1-\theta)$$

- *Expectation is*

$$\mathbb{E}[X] = \frac{r}{\theta}$$

- *Variance is*

$$\mathbb{V}ar(X) = \frac{r(1-\theta)}{\theta^2}$$

5.5 Geometric Distribution

Definition 28. Suppose that independent trials, each having probability of success $0 < \theta < 1$, are performed until a success occurs. If X is the number of trials required, then X is said to have a geometric distribution. It is denoted as

$$X \sim \text{Geo}(\theta)$$

Theorem (Properties of a Geometric Distribution)**.** *For a geometric random variable* $X \sim \text{Geo}(\theta)$

- *The probability distribution of X is*

$$f(x;\theta) = \mathbb{P}(X = x) = \theta(1-\theta)^{x-1} \quad \text{for } x = 1, 2, 3, \ldots$$

- *Moment generating function (mgf) is*

$$M_X(t) = \frac{\theta e^t}{1-(1-\theta)e^t} \qquad \text{for } t < -\ln(1-\theta)$$

- *Expectation is*

$$\mathbb{E}[X] = \frac{1}{\theta}$$

- *Variance is*

$$\mathbb{V}ar(X) = \frac{(1-\theta)}{\theta^2}$$

5.6 Hypergeometric Distribution

To obtain a formula analogous to that of the binomial distribution that applies to sampling without replacement, in which case the trials are not independent, let us consider a set of N elements of which M are looked upon as successes and the other $N - M$ as failures. As in connection with the binomial distribution, we are interested in the probability of getting x successes in n trials, but now we are choosing, without replacement, n of the N elements contained in the set.

Definition 29. A random variable X has a hypergeometric distribution and is referred to as a hypergeometric random variable if and only if its probability distribution is given by

$$f(x; n, N, M) = \mathbb{P}(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

for $x = 0, 1, 2, \ldots, n, \quad x \leq M$ and $n - x \leq N - M$

Theorem (Properties of a Hypergeometric Distribution)**.** *For a hypergeometric random variable*
$X \sim \text{Hypergeo}(n, N, M)$

- *The probability distribution of X is*

$$f(x; n, N, M) = \mathbb{P}(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

  for $x = 0, 1, 2, \ldots, n, \quad x \leq M$ and $n - x \leq N - M$

- *Expectation is*

$$\mu = \frac{nM}{N}$$

- *Variance is*

$$\mathbb{V}ar(X) = \sigma^2 = \frac{nM(N-M)(N-n)}{N^2(N-1)}$$

5.7 Poisson Distribution

Let the discrete random variable X denote the number of times an event occurs in an interval of time (or space). Then X may be a Poisson random variable with $x = 0, 1, 2, \ldots$

Definition 30. A random variable X has a Poisson distribution with parameter $\lambda > 0$ if

$$\mathbb{P}(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}$$

for $x = 0, 1, 2, \ldots$ and $\lambda > 0$. To indicate such random variable, we simply write

$$X \sim \text{Poisson}(\lambda)$$

Theorem (Properties of a Poisson Distribution)**.** *For a Poisson random variable*
$X \sim \text{Poisson}(\lambda)$

- *The probability distribution of X is*

$$\mathbb{P}(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}$$

- *Moment generating function (mgf) is*

$$M_X(t) = e^{\lambda(e^t - 1)}$$

- *Expectation is*
$$\mathbb{E}[X] = \lambda$$
- *Variance is*
$$\mathbb{V}ar(X) = \lambda$$

6 Chapter 6: Special Probability Densities

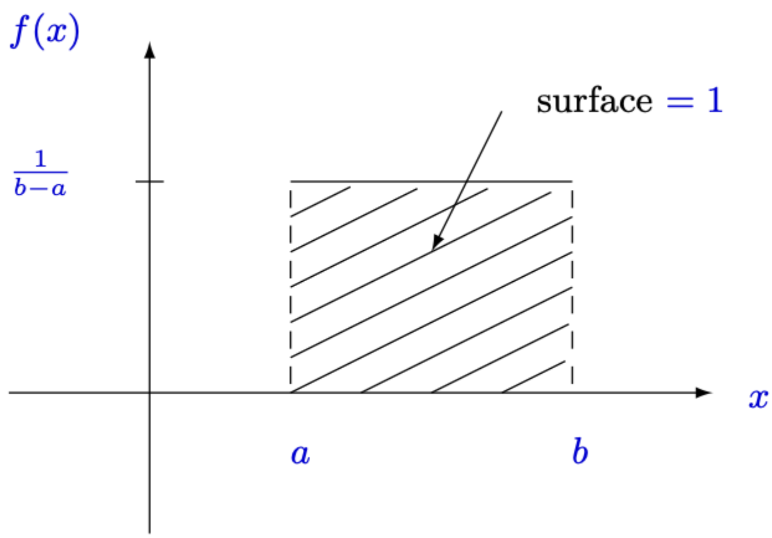
6.1 Uniform Distribution

Definition 31. A random variable X is said to be a uniform random variable over the interval (a, b) if its probability density function is

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

It is denoted as

$$X \sim \text{Unif}(a, b)$$

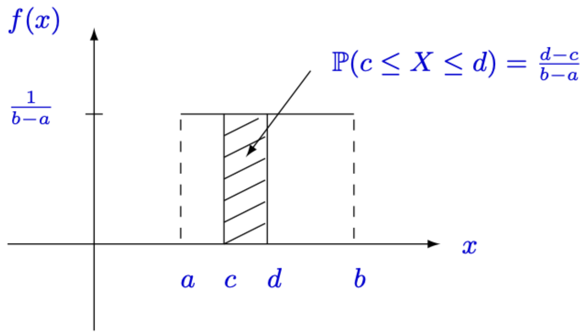


Theorem. *The cumulative distribution function of a uniform random variable X is*

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

If $X \sim \text{Unif}(a, b)$, then:

$$\mathbb{P}(c \leq X \leq d) = \frac{d-c}{b-a} \qquad \forall \ c, d \in [a, b].$$



Theorem (Properties of Uniform Distribution)**.** *For a Uniform random variable* $X \sim \text{Unif}(a, b)$

- *Moment generating function is*

$$M_X(t) = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

- *Expectation is*

$$\mathbb{E}[X] = \frac{(a+b)}{2}$$

- *Variance is*

$$\mathbb{V}ar(X) = \frac{(b-a)^2}{12}$$

6.2 Gamma Function

Definition 32. Let $\alpha > 0$, consider

$$\int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

One can verify that this integral converges if and only if $\alpha > 0$. This integral is known as the Gamma function and denoted by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

Theorem (Properties of the Gamma Function)**.** *The Gamma function satisfies the following properties:*

1. $\Gamma(1) = 1$
2. $\Gamma(2) = 1$
3. $\Gamma(n) = (n-1)!$, for $n \in \mathbb{N}$
4. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \forall \alpha > 0$
5. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Definition 33. A random variable is said to have a gamma distribution with parameters $(\alpha, \beta), \alpha > 0, \beta > 0$, if its density function is given by

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\Gamma(\alpha)$, called the gamma function, is defined as

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

We write $X \sim \text{Gamma}(\alpha, \beta)$

- *Moment generating function (mgf) is*

$$M_X(t) = \frac{1}{(1-\beta t)^\alpha} \quad \text{for } t < \frac{1}{\beta}$$

- *Expectation is*

$$\mathbb{E}[X] = \alpha\beta$$

- *Variance is*

$$\mathbb{V}ar(X) = \alpha\beta^2$$

Theorem. *Let $X \sim Gamma(\alpha, \beta)$, then $f(x; \alpha, \beta)$ is a well defined density, i.e.*

$$\int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = 1$$

6.3 Exponential Distribution

Definition 34. A random variable X has an exponential distribution and it is referred to as an exponential random variable if and only if its probability density is given by

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

It is denoted as

$$X \sim \text{Exp}(\lambda)$$

Theorem (Properties of the Exponential Distribution)**.** *For an exponential random variable* $X \sim \text{Exp}(\lambda)$

- *Moment generating function (mgf) is*

$$M_X(t) = \frac{\lambda}{\lambda - t} \quad \text{for } t < \lambda$$

- *Expectation is*

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

- *Variance is*

$$\mathbb{V}ar(X) = \frac{1}{\lambda^2}$$

6.4 Chi-square Distribution

Definition 35. Let X follow a gamma distribution with $\alpha = \frac{n}{2}$ and $\beta = 2$ (i.e. $X \sim \text{Gamma}(\frac{n}{2}, 2)$), where n is a positive integer. Then the probability density function of X is

$$f(x) = \begin{cases} \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

We say that X follows a chi-square distribution with n degrees of freedom, denoted $X \sim X_n^2$ and read *chi-square-n.*

Theorem (Properties of a Chi-square distribution)**.** *For a Chi-square distribution* $X \sim X_n^2$

- *Moment generating function (mgf) is*

$$M_X(t) = \frac{1}{(1-2t)^{\frac{n}{2}}}$$

- *Expectation is*

$$\mathbb{E}[X] = n$$

- *Variance is*

$$\mathbb{V}ar(X) = 2n$$

6.5 Beta Distribution

Definition 36. A random variable is said to have a beta distribution if its density is given by

$$f(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha, \beta > 0$, and

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$$

is the Beta function. We write $X \sim \text{Beta}(\alpha, \beta)$ to indicate such distribution.

Theorem (Relationship between Beta and Gamma functions)**.** *We have,*

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Theorem (Properties of a Beta distribution)**.** *For a Beta distribution, we have* $X \sim \text{Beta}(\alpha, \beta)$

- $$\mathbb{E}\left[X^k\right] = \frac{B(\alpha+k, \beta)}{B(\alpha, \beta)}, \qquad k > -\alpha$$
- *Moment generating function (mgf) is*
$$M_X(t) = 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{\alpha+r}{\alpha+\beta+r} \right) \frac{t^k}{k!}$$
- *Expectation is*
$$\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$$
- *Variance is*
$$\mathbb{V}ar(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

6.6 The Normal Distribution

Definition 37. The continuous random variable X follows a normal distribution with parameters μ and σ if its probability density function is defined as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\}$$

for $-\infty < x < \infty, -\infty < \mu < \infty, 0 < \sigma < \infty$. We denote $X \sim \mathcal{N}(\mu, \sigma^2)$ to indicate such distribution.

Theorem (Properties of a Normal distribution). *Let $X \sim \mathcal{N}(\mu, \sigma^2)$, then*

- *A well defined pdf f_X is*

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(x-\mu)^2}{2\sigma^2} \right) dx = 1$$

- *Expectation is*

$$\mathbb{E}[X] = \mu$$

- *Variance is*

$$\mathbb{V}ar(X) = \sigma^2$$

Theorem. *If $X \sim \mathcal{N}(\mu, \sigma^2)$, then*

$$Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

In particular, if $Z = \frac{X-\mu}{\sigma}$, then $Z \sim \mathcal{N}(0,1)$. The moment generating functions for each are

- *For $Z \sim \mathcal{N}(0,1)$*

$$M_Z(t) = \exp\left(\frac{t^2}{2}\right)$$

- *For $X \sim \mathcal{N}(\mu, \sigma^2)$*

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

6.7 Normal Distribution to Other Distributions

Theorem (Normal distribution to Chi-square distribution). *If X is normally distributed with mean μ and variance $\sigma^2 > 0$, then*

$$U = \left(\frac{X-\mu}{\sigma}\right)^2 = Z^2$$

is distributed as a chi-square random variable with 1 degree of freedom.

Theorem (Normal approximation to Binomial distribution). *If X is a random variable having a binomial distribution with the parameters n and θ, then*

$$Z = \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}}$$

approaches that of the standard normal distribution when $n \to \infty$.

6.8 Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
+0.1	.53983	.54380	.54776	.55172	.55567	.55966	.56360	.56749	.57142	.57535
+0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
+0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
+0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
+0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
+0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
+0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
+0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
+0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
+1	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
+1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
+1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
+1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91308	.91466	.91621	.91774
+1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
+1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
+1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
+1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
+1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
+2	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
+2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
+2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
+2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
+2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
+2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
+2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
+2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
+2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
+2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
+3	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
+3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
+3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
+3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
+3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
+3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
+3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
+3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
+3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
+3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997
+4	.99997	.99997	.99997	.99997	.99997	.99997	.99998	.99998	.99998	.99998

Figure 1: Positive Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
-0.1	.46017	.45620	.45224	.44828	.44433	.44034	.43640	.43251	.42858	.42465
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-1	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-2	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-3	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-4	.00003	.00003	.00003	.00003	.00003	.00003	.00002	.00002	.00002	.00002

Figure 2: Negative Standard Normal Distribution Table