Math 461 Spring 2024

Renming Song

University of Illinois Urbana-Champaign

February 16, 2024

Outline

Outline

HW4 is due today, before the end of class.

I will post solutions to HW4 this afternoon.

HW4 is due today, before the end of class.

I will post solutions to HW4 this afternoon.

Outline

Example 5

n balls are randomly distributed into r boxes (so that each ball is equally likely to go to any of the r boxes). Find the expected number of empty boxes.

Let X be the number of boxes. For $i = 1, \ldots, r$, let $X_i = 1$ if box number i is empty and $X_i = 0$ otherwise. Then $X = X_1 + \cdots + X_r$. Note that for $i = 1, \ldots, r$,

$$E[X_i] = P(X_i = 1) = \left(\frac{r-1}{r}\right)^n.$$

Thus

$$E[X] = E[X_1] + \cdots + E[X_r] = r\left(\frac{r-1}{r}\right)^n.$$

Example 5

n balls are randomly distributed into r boxes (so that each ball is equally likely to go to any of the r boxes). Find the expected number of empty boxes.

Let X be the number of boxes. For $i = 1, \ldots, r$, let $X_i = 1$ if box number i is empty and $X_i = 0$ otherwise. Then $X = X_1 + \cdots + X_r$. Note that for $i = 1, \ldots, r$,

$$E[X_i] = P(X_i = 1) = \left(\frac{r-1}{r}\right)^n.$$

Thus

$$E[X] = E[X_1] + \cdots + E[X_r] = r\left(\frac{r-1}{r}\right)^n.$$

Example 6

Suppose that there are N types of coupons and each time one gets a coupon it is equally likely to be any one of the N types. (a) If you have just collected n coupons, what is the expected number of different types of coupons in your collection? (b) Find the expected number of coupons you need to amass in order to get a complete set.

Let X be the number of different types in your collection of n coupons. For $i = 1, \ldots, N$, let $X_i = 1$ if there is at least one type i coupon in your collection and $X_i = 0$ otherwise. Then $X = X_1 + \cdots + X_N$. Note that, for $i = 1, \ldots, N$,

$$E[X_i] = P(X_i = 1) = 1 - \left(\frac{N-1}{N}\right)^n.$$

Thus

$$E[X] = E[X_1] + \cdots + E[X_N] = N\left(1 - \left(\frac{N-1}{N}\right)^n\right).$$

Example 6

Suppose that there are N types of coupons and each time one gets a coupon it is equally likely to be any one of the N types. (a) If you have just collected n coupons, what is the expected number of different types of coupons in your collection? (b) Find the expected number of coupons you need to amass in order to get a complete set.

Let X be the number of different types in your collection of n coupons. For $i = 1, \ldots, N$, let $X_i = 1$ if there is at least one type i coupon in your collection and $X_i = 0$ otherwise. Then $X = X_1 + \cdots + X_N$. Note that, for $i = 1, \ldots, N$,

$$E[X_i] = P(X_i = 1) = 1 - \left(\frac{N-1}{N}\right)^n.$$

Thus

$$E[X] = E[X_1] + \cdots + E[X_N] = N\left(1 - \left(\frac{N-1}{N}\right)^n\right).$$

(b) Let Y be the number of coupons you need to amass in order to get a complete set. Let $Y_1 = 1$; let Y_2 be the number of additional coupons needed in order to get a new type (i.e., different from the one you got); let Y_3 be the number of additional coupons, after you have got two types, needed in order to get a new type (different from the two types you already got); $\dots$, let Y_N be the number of additional coupons, after you have got $N-1$ types, to get the final type. Then $Y = Y_1 + \cdots + Y_N$.

Y_2 is a geometric random variable with parameter $\frac{N-1}{N}$; Y_3 is a geometric random variable with parameter $\frac{N-2}{N}$, $\dots$, Y_N is a geometric random variable with parameter $\frac{1}{N}$. Thus

$$E[Y] = E[Y_1] + \cdots + E[Y_N] = 1 + \frac{N}{N-1} + \cdots + N = N\left(1 + \frac{1}{2} + \cdots + \frac{1}{N}\right).$$

(b) Let Y be the number of coupons you need to amass in order to get a complete set. Let $Y_1 = 1$; let Y_2 be the number of additional coupons needed in order to get a new type (i.e., different from the one you got); let Y_3 be the number of additional coupons, after you have got two types, needed in order to get a new type (different from the two types you already got); ..., let Y_N be the number of additional coupons, after you have got $N-1$ types, to get the final type. Then $Y = Y_1 + \cdots + Y_N$.

Y_2 is a geometric random variable with parameter $\frac{N-1}{N}$; Y_3 is a geometric random variable with parameter $\frac{N-2}{N}$, ..., Y_N is a geometric random variable with parameter $\frac{1}{N}$. Thus

$$E[Y] = E[Y_1] + \cdots + E[Y_N] = 1 + \frac{N}{N-1} + \cdots + N = N(1 + \frac{1}{2} + \cdots + \frac{1}{N}).$$

Outline

The distribution function F of a random variable X is defined by

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

The distribution function F of any random variable X satisfies

(i) F is non-decreasing;

(ii) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$;

(iii) F is right-continuous, i.e., $F(x+) = F(x)$ for every $x \in \mathbb{R}$.

We know that

$$P(X < b) = F(b-).$$

The distribution function F of a random variable X is defined by

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

The distribution function F of any random variable X satisfies

(i) F is non-decreasing;

(ii) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$;

(iii) F is right-continuous, i.e., $F(x+) = F(x)$ for every $x \in \mathbb{R}$.

We know that

$$P(X < b) = F(b-).$$

The distribution function F of a random variable X is defined by

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

The distribution function F of any random variable X satisfies

(i) F is non-decreasing;

(ii) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$;

(iii) F is right-continuous, i.e., $F(x+) = F(x)$ for every $x \in \mathbb{R}$.

We know that

$$P(X < b) = F(b-).$$

In general, the distribution function F of a random variable is not continuous.

A random variable X is said to be continuous if its distribution function is continuous, or equivalently,

$$P(X = x) = 0, \quad \text{for every } x \in \mathbb{R}.$$

If X is a continuous random variable, then for any $a \in \mathbb{R}$,

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

So computations will be a lot easier for continuous random variables.

In general, the distribution function F of a random variable is not continuous.

A random variable X is said to be continuous if its distribution function is continuous, or equivalently,

$$P(X = x) = 0, \quad \text{for every } x \in \mathbb{R}.$$

If X is a continuous random variable, then for any $a \in \mathbb{R}$,

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

So computations will be a lot easier for continuous random variables.

In general, the distribution function F of a random variable is not continuous.

A random variable X is said to be continuous if its distribution function is continuous, or equivalently,

$$P(X = x) = 0, \quad \text{for every } x \in \mathbb{R}.$$

If X is a continuous random variable, then for any $a \in \mathbb{R}$,

$$P(X < a) = P(X \le a).$$

So computations will be a lot easier for continuous random variables.

General random variables can be classified as discrete random variables; continuous random variables; or neither discrete nor continuous.

The random variables with distribution function

$$F(x) = \begin{cases} 0, & x < 0, \\ x/3, & 0 \le x < 1, \\ x/2, & 1 \le x < 2, \\ 1, & x \ge 2. \end{cases}$$

is neither discrete nor continuous.

In Chapter 4, we learned about discrete random variables. In this chapter, we will concentrate on continuous random variables.

General random variables can be classified as discrete random variables; continuous random variables; or neither discrete nor continuous.

The random variables with distribution function

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

is neither discrete nor continuous.

In Chapter 4, we learned about discrete random variables. In this chapter, we will concentrate on continuous random variables.

General random variables can be classified as discrete random variables; continuous random variables; or neither discrete nor continuous.

The random variables with distribution function

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

is neither discrete nor continuous.

In Chapter 4, we learned about discrete random variables. In this chapter, we will concentrate on continuous random variables.

In fact, we will concentrate on a subclass of continuous random variables – the so-called absolutely continuous random variables.

Definition

A non-negative function f on $\mathbb{R}$ is called a probability density function if

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

If f is a probability density function, then the function F defined by

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R} \tag{1}$$

is a continuous distribution function.

In fact, we will concentrate on a subclass of continuous random variables – the so-called absolutely continuous random variables.

Definition

A non-negative function f on $\mathbb{R}$ is called a probability density function if

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

If f is a probability density function, then the function F defined by

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R} \tag{1}$$

is a continuous distribution function.

In fact, we will concentrate on a subclass of continuous random variables – the so-called absolutely continuous random variables.

Definition

A non-negative function f on $\mathbb{R}$ is called a probability density function if

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

If f is a probability density function, then the function F defined by

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R} \tag{1}$$

is a continuous distribution function.

Definition

A random variable X is said to be absolutely continuous if there is a non-negative function f on $\mathbb{R}$ such that

$$P(X \leq x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}.$$

f must be a probability density and it is called the density of X.

The density function of an absolutely continuous random variable X contains all the statistical info about X. If we know the density f, we can get the distribution F of X via (1). If we know the the distribution F of an absolutely continuous random variable X, we can simply differentiate F to get the density f. For points where F is not differentiable, we simply let f equal 0 there.

Definition

A random variable X is said to be absolutely continuous if there is a non-negative function f on $\mathbb{R}$ such that

$$P(X \leq x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}.$$

f must be a probability density and it is called the density of X.

The density function of an absolutely continuous random variable X contains all the statistical info about X. If we know the density f, we can get the distribution F of X via (1). If we know the the distribution F of an absolutely continuous random variable X, we can simply differentiate F to get the density f. For points where F is not differentiable, we simply let f equal 0 there.

If we know the density f of an absolutely continuous random variable X, then

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

There exists continuous random variables that are not absolutely continuous. But we are not going to deal with these type of "pathological" random variables in this course.

If we know the density f of an absolutely continuous random variable X, then

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

There exists continuous random variables that are not absolutely continuous. But we are not going to deal with these type of "pathological" random variables in this course.

Example 1

Choose a point at random from the disk $B(0,1)$. Let X be the distance between the chosen point and the origin.

For any $x \in [0,1]$,

$$P(X \leq x) = \frac{\pi x^2}{\pi} = x^2.$$

Thus the distribution function of X is

$$F(x) = \begin{cases} 0, & x \leq 0, \\ x^2, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

The density of X is

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 1

Choose a point at random from the disk $B(0,1)$. Let X be the distance between the chosen point and the origin.

For any $x \in [0,1]$,

$$P(X \leq x) = \frac{\pi x^2}{\pi} = x^2.$$

Thus the distribution function of X is

$$F(x) = \begin{cases} 0, & x \leq 0, \\ x^2, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

The density of X is

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$