Section 4/21

Tuesday, April 21, 2020 10:36 AM

Announcements:

- Quiz:
  - sent via Bcourses message at **12 pm PST today**.
  - open book, on sections 12.7 and 8.4.
  - due at **12 pm PST Weds (tomorrow).**
  - Turn in solutions on **4 sheets, 1 problem per page as 1 PDF document.** To encourage you to figure out how to format your submission, **I will deduct 2 pts** if your submission is not formatted appropriately.

Continuous Probability (10.1)

Probability is the study of **random variables** and **probability functions.**

As an example, say you have a coin with the label "0" on one side and "1" on the other side. Then the 0/1 value of the coin when flipped is a random variable. A fair coin has 1/2 change of being heads and 1/2 chance of being tails, so the probability function is

$$f : \{0,1\} \to \mathbb{R} \qquad f(0) = \frac{1}{2} \text{ and } f(1) = \frac{1}{2}$$

Definition 1. A **discrete probability function** on some set of numbers $\{x_1, x_2, \dots, x_k\}$ (called **events**) is an assignment of a **probability**

$$f(x_i) \text{ to each event } x_i$$

satisfying some obvious constraints.

a) The probability of something happening is non-negative.

$$0 \leq f(x_i)$$

b) The sum of the probabilities is 1: something has to happen.

$$\sum_{i=1}^{k} f(x_i) = 1$$

Often you are in a situation where the events are not a finite set, but rather a real number in some interval. A **continuous random variable** X is a random variable that takes values in some interval $[a, b]$.

Definition 2. The **probability function** for a continuous random variable X, with values in the interval $[a, b]$, is a function of the form

$f: [a, b] \to \mathbb{R}$ from $[a, b]$ to the real numbers.

The value $f(x)$ can be used to calculate

a) The probability of a particular event is non-negative.

$f(x) \geq 0$ for any $a \leq x \leq b$

b) The total probability is 1. In the continuous case, this is stated using an integral.

$$\int_a^b f(x)dx = 1$$

The **probability** $P(c \leq X \leq d)$ is given by the integral

$$P(c \leq X \leq d) = \int_c^d f(x)dx$$

and the **cumulative distribution** $F(x)$ is the function given by

$$F(x) = P(a \leq X) = \int_a^x f(y)dy$$

Example 1. Let $f\colon [a,b] \to \mathbb{R}$ be the constant function $f(x) = \frac{1}{b-a}$. Let's check that this is a probability distribution.

(a) holds since $f(x) = \frac{1}{b-a} \geq 0$ (b) holds since $\int_a^b \frac{1}{a-b} dx = \frac{b-a}{b-a} = 1$

This is called the **uniform** distribution, since every event x has the same probability. The probability $P(c \leq X \leq d)$ and cumulative distribution $F(x)$ are given by

$$P(c \leq X \leq d) = \int_c^d \frac{1}{a-b} dx = \frac{d-c}{b-a} \qquad F(x) = = \frac{x-c}{b-a}$$

Exercises For You. Here are some exercises for you to try.

1) Is $f(x) = 12x^3$ a probability distribution on the interval $[1,2]$?

2) The battery in the Voyager spacecraft will randomly stop function at some point in the future after the time it was launched ($t = 0$ in years). The probability of failure is given by

$$f(t) = ke^{-t/2}$$

(a) Find the value of k that makes this into a probability distribution function on $[0, \infty)$. (b) Find the probability that the battery will fail within 5 years.

Expected Value And Standard Deviation (10.2).

There are some key quantities that can be associated to a random variable: the *mean*, which calculates the average value that a random variable will take, and the *standard deviation*, which measures how much the random variable typically deviates from the mean.

Now we'll give the actual definitions. Let's start with the discrete version.

Definition 3. Let f be a discrete probability distribution on $\{x_1, x_2, \ldots, x_k\}$.

- The **expected value** $E(X)$ of X, also called the **mean** and denoted by μ, is defined by the formula.

$$E(X) = \sum_{i=1}^{k} x_i f(x_i)$$

  That is, it's the sum over each event of the *probability* times the *value.*

- The **variance** $Var(X)$ of X, is defind by the formula

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 f(x_i)$$

- The **standard deviation** $StDev(X)$, also denoted by σ, is denoted by

$$StDev(X) = \sqrt{Var(X)}$$

Next we can give the continuous version. In this case, we can also define the *meidian.*

Definition 4. Let X be a random variable with probability distribution f on the interval $[a, b]$.

- The **expected value** $E(X)$ of X (or **mean** μ), is the integral

$$E(X) = \int_a^b xf(x)dx$$

- The **variance** $Var(X)$ of X, is defind by the formula

$$Var(X) = \int_a^b (x-\mu)^2 f(x)dx$$

- The **standard deviation** $StDev(X)$, also denoted by σ, is denoted by

$$StDev(X) = \sqrt{Var(X)}$$

- The **median** $Med(X) = m$ is the unique number m with $a \leq m \leq b$ that satisfies the following identity.

$$\int_a^m f(x)dx = F(m) = 1/2$$

A useful fact is that the variance can be calculated via another formula.

Formula 1. The variance $Var(X)$ can also be given by the formula

$$Var(X) = \left(\int_a^b x^2 f(x)dx\right) - \mu^2$$

Example 2. Let X be a random variable with probability distribution $f: [1,2] \to \mathbb{R}$ given by the formula

$$f(x) = \frac{2}{3}x$$

Let's calculate the mean, variance and standard deviation. The expected value is given by

$$E(X) = \int_1^2 \frac{2}{3}x^2 dx = \frac{2}{9}2^3 - \frac{2}{9}1^3 = \frac{14}{9} \approx 1.555$$

The variance can now be calculated using either the definition or Formula 1. Using Formula 1, we find that:

$$Var(X) = \int_1^2 \frac{2}{3}x^3 dx - \left(\frac{14}{9}\right)^2 = \left(\frac{2}{12}2^4 - \frac{2}{12}1\right) - \left(\frac{14}{9}\right)^2 \approx .08$$

Finally, the standard deviation is $StDev(X) \approx .28$

Exercises For You. Now you should try these calculations.

3) Recall the failure probability distribution for the battery in the Voyager spacecraft

$$f(t) = ke^{-t/2} = \frac{1}{2}e^{-t/2}$$

Compute the expected failure time and the standard deviation of the failure time of the battery.

4) Is the median of a continuous random variable always less than the mean? Is the mean always less than the median? Why or why not?

Section 4/21

Tuesday, April 21, 2020
10:36 AM

Announcements:

- Quiz:
  - sent via Bcourses message at **12 pm PST today**.
  - open book, on sections 12.7 and 8.4.

- due at **12 pm PST Weds (tomorrow).**
- Turn in solutions on **4 sheets, 1 problem per page as 1 PDF document.** To encourage you to figure out how to format your submission, **I will deduct 2 pts** if your submission is not formatted appropriately.

Continuous Probability (10.1)

Probability is the study of **random variables** and **probability functions.**

As an example, say you have a coin with the label "0" on one side and "1" on the other side. Then the 0/1 value of the coin when flipped is a random variable. A fair coin has 1/2 change of being heads and 1/2 chance of being tails, so the probability function is

$$f: \{0,1\} \to \mathbb{R} \qquad f(0) = \frac{1}{2} \text{ and } f(1) = \frac{1}{2}$$

Definition 1. A **discrete probability function** on some set of numbers $\{x_1, x_2, \dots, x_k\}$ (called **events**) is an assignment of a **probability**

$f(x_i)$ to each event x_i

satisfying some obvious constraints.

a. The probability of something happening is non-negative.

$$0 \le f(x_i)$$

a. The sum of the probabilities is 1: something has to happen.

$$\sum_{i=1}^{k} f(x_i) = 1$$

Often you are in a situation where the events are not a finite set, but rather a real number in some interval. A **continuous random variable** X is a random variable that takes values in some interval $[a, b]$.

Definition 2. The **probability function** for a continuous random variable X, with values in the interval $[a, b]$, is a function of the form

$$f: [a, b] \to \mathbb{R} \quad \text{from } [a, b] \text{ to the real numbers.}$$

The value $f(x)$ can be used to calculate

a. The probability of a particular event is non-negative.

$$f(x) \geq 0 \quad \text{for any } a \leq x \leq b$$

a. The total probability is 1. In the continuous case, this is stated using an integral.

$$\int_a^b f(x)dx = 1$$

The **probability** $P(c \leq X \leq d)$ is given by the integral

$$P(c \leq X \leq d) = \int_c^d f(x)dx$$

and the **cumulative distribution** $F(x)$ is the function given by

$$F(x) = P(a \leq X) = \int_a^x f(y)dy$$

Example 1. Let $f: [a, b] \to \mathbb{R}$ be the constant function $f(x) = \frac{1}{b-a}$.

Let's check that this is a probability distribution.

(a) holds since $f(x) = \frac{1}{b-a} \geq 0$ (b) holds since $\int_a^b \frac{1}{a-b} dx = \frac{b-a}{b-a} = 1$

This is called the **uniform** distribution, since every event x has the same probability. The probability $P(c \leq X \leq d)$ and cumulative distribution $F(x)$ are given by

$$P(c \leq X \leq d) = \int_c^d \frac{1}{a-b} dx = \frac{d-c}{b-a} \qquad F(x) == \frac{x-c}{b-a}$$

Exercises For You. Here are some exercises for you to try.

1. Is $f(x) = 12x^3$ a probability distribution on the interval $[1,2]$?

2. The battery in the Voyager spacecraft will randomly stop function at some point in the future after the time it was launched ($t = 0$ in years). The probability of failure is given by

$$f(t) = ke^{-t/2}$$

(a) Find the value of k that makes this into a probability distribution function on $[0, \infty)$. (b) Find the probability that the battery will fail within 5 years.

__Expected Value And Standard Deviation (10.2).__

There are some key quantities that can be associated to a random variable: the *mean*, which calculates the average value that a random variable will take, and the *standard deviation*, which measures how much the random variable typically deviates from the mean.

Now we'll give the actual definitions. Let's start with the discrete version.

Definition 3. Let f be a discrete probability distribution on $\{x_1, x_2, \dots, x_k\}$.

- The **expected value** $E(X)$ of X, also called the **mean** and denoted by μ, is defined by the formula.

$$E(X) = \sum_{i=1}^{k} x_i f(x_i)$$

That is, it's the sum over each event of the *probability* times the *value.*

- The **variance** $Var(X)$ of X, is defind by the formula

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 f(x_i)$$

- The **standard deviation** $StDev(X)$, also denoted by σ, is denoted by

$$StDev(X) = \sqrt{Var(X)}$$

Next we can give the continuous version. In this case, we can also define the *meidian.*

Definition 4. Let X be a random variable with probability distribution f on the interval $[a, b]$.

- The **expected value** $E(X)$ of X (or **mean** μ), is the integral

$$E(X) = \int_a^b x f(x) dx$$

- The **variance** $Var(X)$ of X, is defind by the formula

$$Var(X) = \int_a^b (x-\mu)^2 f(x)dx$$

- The **standard deviation** $StDev(X)$, also denoted by σ, is denoted by

$$StDev(X) = \sqrt{Var(X)}$$

- The **median** $Med(X) = m$ is the unique number m with $a \leq m \leq b$ that satisfies the following identity.

$$\int_a^m f(x)dx = F(m) = 1/2$$

A useful fact is that the variance can be calculated via another formula.

Formula 1. The variance $Var(X)$ can also be given by the formula

$$Var(X) = \left(\int_a^b x^2 f(x)dx\right) - \mu^2$$

Example 2. Let X be a random variable with probability distribution $f: [1,2] \to \mathbb{R}$ given by the formula

$$f(x) = \frac{2}{3}x$$

Let's calculate the mean, variance and standard deviation. The expected value is given by

$$E(X) = \int_1^2 \frac{2}{3}x^2 dx = \frac{2}{9}2^3 - \frac{2}{9}1^3 = \frac{14}{9} \approx 1.555$$

The variance can now be calculated using either the definition or

Formula 1. Using Formula 1, we find that:

$$Var(X) = \int_1^2 \frac{2}{3} x^3 dx - \left(\frac{14}{9}\right)^2 = \left(\frac{2}{12} 2^4 - \frac{2}{12} 1\right) - \left(\frac{14}{9}\right)^2 \approx .08$$

Finally, the standard deviation is $StDev(X) \approx .28$

Exercises For You. Now you should try these calculations.

1. Recall the failure probability distribution for the battery in the Voyager spacecraft

$$f(t) = ke^{-t/2} = \frac{1}{2} e^{-t/2}$$

Compute the expected failure time and the standard deviation of the
failure time of the battery.

1. Is the median of a continuous random variable always less than the mean? Is the mean always less than the median? Why or why not?