Density and Probability Functions

Gene Quinn

Assigning Probabilities

For a continuous random variable X, we are usually interested in events of the form "X is in the interval $[a, b]$"

Assigning Probabilities

For a continuous random variable X, we are usually interested in events of the form "X is in the interval $[a, b]$"

The mechanism for assigning probabilities to events of this type,

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

is called a **probability density function** (PDF), denoted by $f(x)$, which is defined to have the property that:

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

Assigning Probabilities

To be valid, a density function has to have the following properties:

$$f(x) \geq 0 \quad \text{for all} \quad x$$

and

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

Assigning Probabilities

To be valid, a density function has to have the following properties:

$$f(x) \geq 0 \quad \text{for all} \quad x$$

and

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

Any function that satisfies these two properties can be thought of as a density function of a random variable.

Assigning Probabilities

To be valid, a density function has to have the following properties:

$$f(x) \geq 0 \quad \text{for all} \quad x$$

and

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

Any function that satisfies these two properties can be thought of as a density function of a random variable.

Some random variables only take values on a subset of the real line, say $[0, \infty)$. In this case, the second condition is

$$\int_{0}^{\infty} f(x)\, dx = 1$$

The Uniform Distribution

The function

$$f(x) = 1 \quad 0 < x < 1$$

is a valid density function, since it is never negative and

$$\int_0^1 f(x)\,dx = x]_0^1 = 1$$

(in fact, it is the density function of an important continuous distribution called the **uniform** distribution)

The Uniform Distribution

The function

$$f(x) = 1 \quad 0 < x < 1$$

is a valid density function, since it is never negative and

$$\int_0^1 f(x)\, dx = x]_0^1 = 1$$

(in fact, it is the density function of an important continuous distribution called the **uniform** distribution)

We can find the probability that X falls into the interval $[.25, .75]$ from the definition as

$$P(.25 \leq X \leq .75) = \int_{.25}^{.75} 1 \cdot dx = x]_{.25}^{.75} = 0.5$$

The Uniform Distribution

The `RAND` function produces a single observation from the uniform distribution.

The Uniform Distribution

The `RAND` function produces a single observation from the uniform distribution.

The R command `x=runif(1000)` generates 1,000 observations from the uniform distribution and stores them in an array called `x`.

The Uniform Distribution

The `RAND` function produces a single observation from the uniform distribution.

The R command `x=runif(1000)` generates 1,000 observations from the uniform distribution and stores them in an array called `x`.

The uniform distribution plays an important role in simulation and in modern statistical procedures that rely on simulating random variables from a specified distribution.

The Cumulative Distribution Function

An important function called the **Cumulative Distribution Function** (CDF) of a continuous random variable X is defined by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$$

where f is the density function of X and t is a dummy variable in the integration

The Cumulative Distribution Function

An important function called the **Cumulative Distribution Function** (CDF) of a continuous random variable X is defined by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$$

where f is the density function of X and t is a dummy variable in the integration

The CDF is usually designated by the uppercase letter corresponding to the lower case letter used to represent the density function.

The Cumulative Distribution Function

An important function called the **Cumulative Distribution Function** (CDF) of a continuous random variable X is defined by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$$

where f is the density function of X and t is a dummy variable in the integration

The CDF is usually designated by the uppercase letter corresponding to the lower case letter used to represent the density function.

Every cumulative distribution function is non-decreasing and approaches zero as $x \rightarrow -\infty$ and one as $x \rightarrow \infty$.