Continuous Random Variables

Section 13

Continuous Random Variables

Definition:

A random variable that takes on a *continuum* of values is called a continuous random variable.

Continuous Random Variables

Example:

Distributions of Lengths of Boa Constrictors

The boa constrictor is a large species of snake that can grow to anywhere between 1 m and 4 m in length.

Let *L* be the continuous random variable that measures the length of a snake.

$$L : S \rightarrow [1,4]$$

Continuous Random Variables

The lengths of 500 boas are recorded below:

Table 13.1

Length range (m)	Frequency	Relative frequency
$[1, 1.5)$	20	0.04
$[1.5, 2)$	58	0.116
$[2, 2.5)$	122	0.244
$[2.5, 3)$	180	0.36
$[3, 3.5)$	86	0.172
$[3.5, 4)$	34	0.068

Note: relative frequency = frequency/500 = probability

Continuous Random Variables

<u>Histogram for Probability Mass</u>:

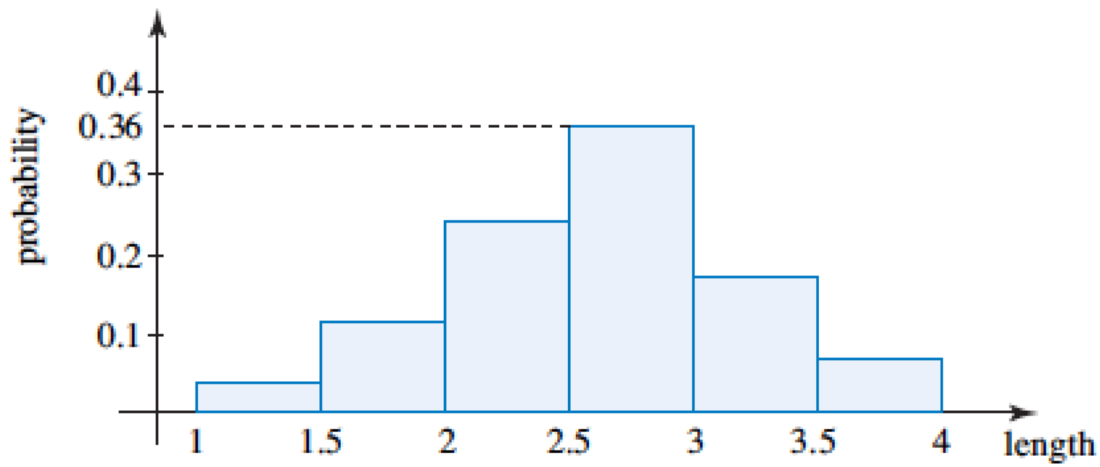


FIGURE 13.1

Histogram: the heights represent the probability

The probability that a randomly selected boa is between 2.5 m and 3 m in length is the <u>height</u> of the rectangle over [2.5, 3), i.e., 0.36.

Continuous Random Variables

To draw a histogram representing probability **density**, we re-label the vertical axis so that the probability that *L* belongs to an interval is the area of the rectangle above that interval.

Continuous Random Variables

Note:

$$\frac{\text{probability } \textbf{mass}}{\textbf{length} \text{ of interval}} = \text{probability } \textbf{density}$$

Continuous Random Variables

For example, consider the interval [2.5, 3). The probability that *L* falls in this range is 0.36.

Now, we want this value to be the area of the rectangle over [2.5, 3), so

probability density (height) = 0.36/(3-2.5) = 0.72

Continuous Random Variables

Histogram for Probability Density:

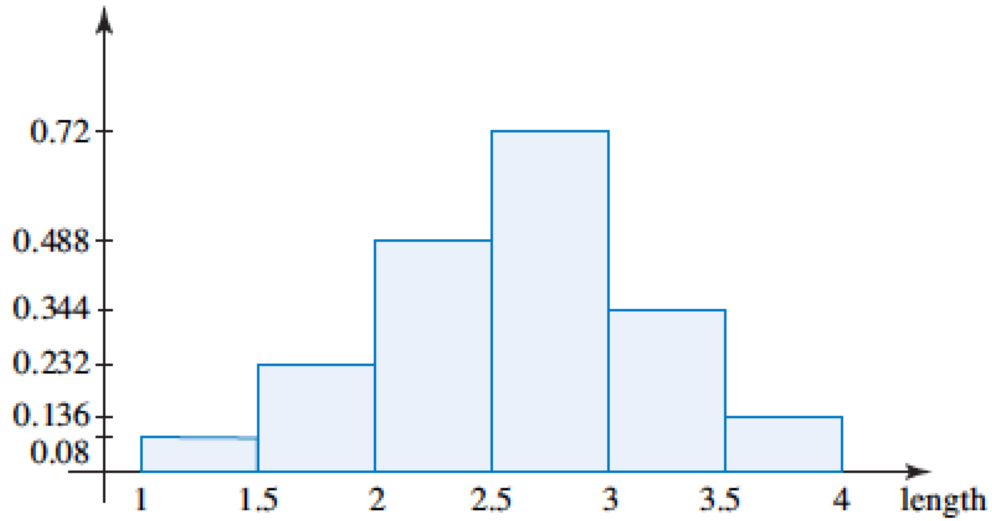


FIGURE 13.2

Histogram: the areas represent the probability

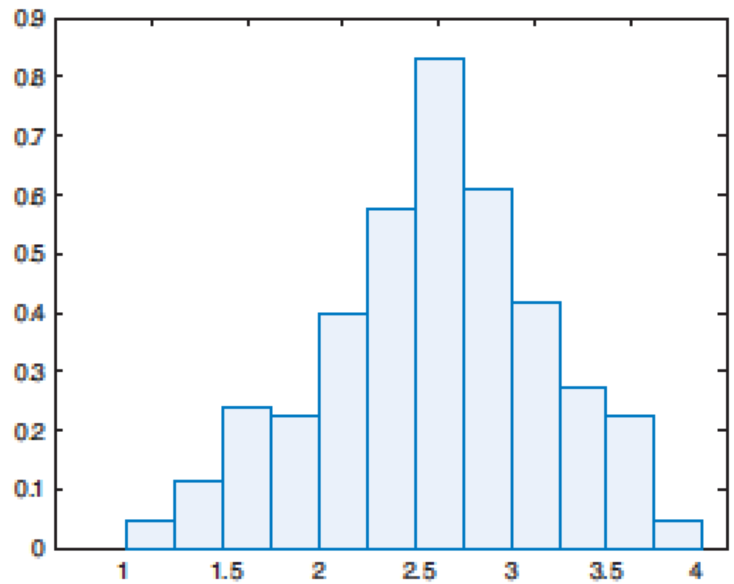
The probability that a randomly selected boa is between 2.5 m and 3 m in length is the area of the rectangle above [2.5, 3), i.e. 0.36.

Continuous Random Variables

To get a more precise probability mass (or density) function, we divide [1,4] into smaller subintervals:

Table 13.2

Length range (m)	Frequency	Relative frequency
$[1, 1.25)$	6	0.012
$[1.25, 1.5)$	14	0.028
$[1.5, 1.75)$	30	0.06
$[1.75, 2)$	28	0.056
$[2, 2.25)$	50	0.1
$[2.25, 2.5)$	72	0.144
$[2.5, 2.75)$	104	0.208
$[2.75, 3)$	76	0.152
$[3, 3.25)$	52	0.104
$[3.25, 3.5)$	34	0.068
$[3.5, 3.75)$	28	0.056
$[3.75, 4)$	6	0.012



Continuous Random Variables

As we continue to increase the number of subintervals, we obtain a more and more refined histogram.

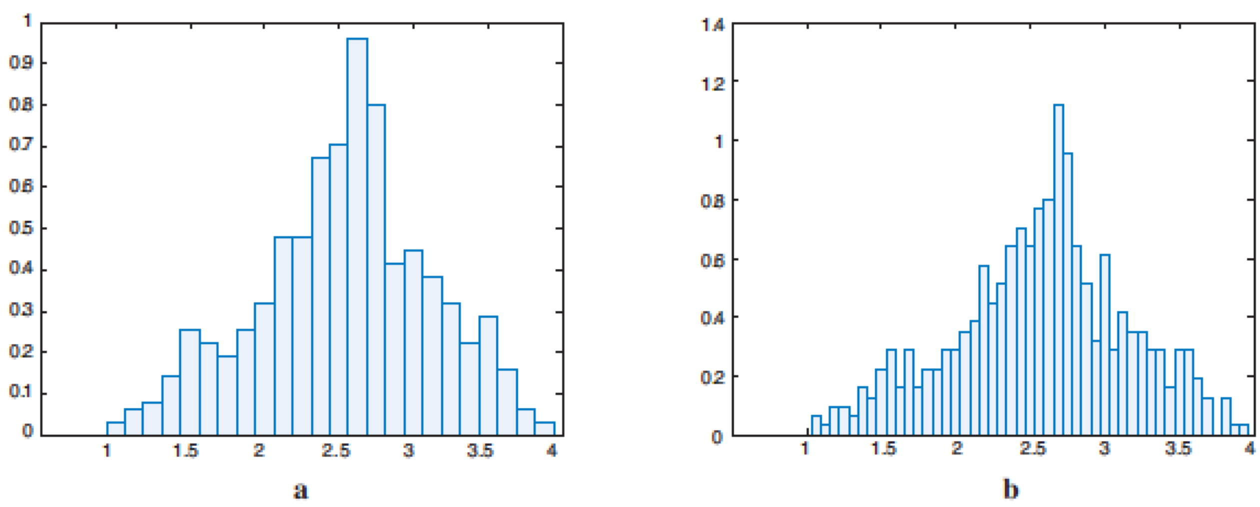


FIGURE 13.4

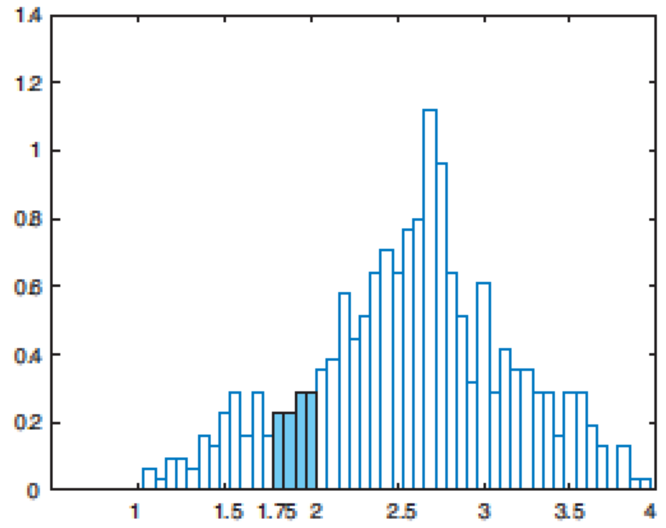
Histograms based on 24 and 48 subintervals

Continuous Random Variables

<u>Riemann Sum</u>:

FIGURE 13.5

Probability of boa length between 1.75 m and 2 m



The probability that a randomly chosen boa is between 1.75 m and 2 m in length is the sum of the areas of the rectangles over the interval [1.75, 2).

Continuous Random Variables

To obtain the probability density **function**, we let the length of the intervals approach 0 and the number of rectangles approach ∞.

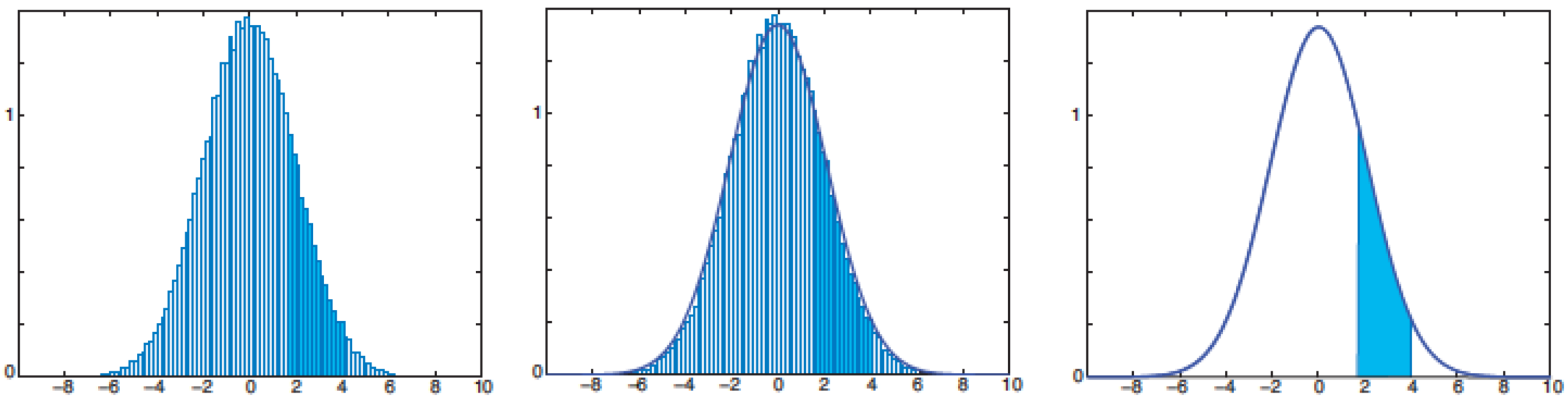


FIGURE 13.6

From a histogram to a density function

Probability Density Functions

Definition: Defining Properties of a PDF
Assume that the interval *I* represents the range of a continuous random variable *X*. A function *f(x)* can be a probability density function if

(1) $f(x) \geq 0 \quad \text{for all } x \in I.$

(2) $\int_I f(x)dx = 1.$

Probability Density Functions

Example:

Show that $f(x) = \dfrac{2}{\pi(1+x^2)}$

could be a probability density function for some continuous random variable on [0, ∞).

Calculating Probabilities

For a continuous random variable, we calculate the probability that a random variable belongs to an *interval* of real numbers.

The probability that an outcome *X* is between *a* and *b* is the area under the graph of *f(x)* on [*a*,*b*]:

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

Calculating Probabilities

The probability that an outcome is *equal* to a particular value is zero.

$$P(a \le X \le a) = \int_a^a f(x)dx = 0$$

For this reason, **including or excluding the endpoints of an interval does not affect the probability**, i.e.,

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

Calculating Probabilities

Example #32:

The distance between a seed and the plant it came from is modelled by the density function

$$f(x) = \frac{2}{\pi(1+x^2)} \qquad x \in [0,\infty).$$

where *x* represents the distance (in metres),

What is the probability that a seed will be found farther than 5 m from the plant?

Cumulative Distribution Function

Definition:

Suppose that *f(x)* is a probability density function defined on an interval [*a,b*]. The function *F(x)* defined by

$$F(x) = P(X \leq x) = \int_a^x f(t)dt$$

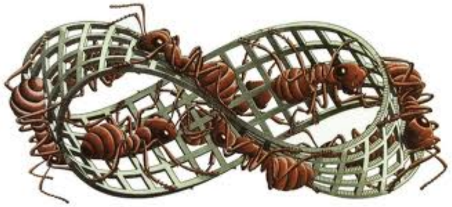
for all *x* in [*a,b*] is called a cumulative distribution function of *f(x)*.

Cumulative Distribution Function

Example #30 (modified):

Suppose that the lifetime of an insect is given by the probability density function

$$f(t) = 0.2e^{-0.2t} \qquad t \in [0,\infty).$$



where *t* is measured in days,

(a) Determine the corresponding cumulative distribution function, *F(t)*.
(b) Find the probability that the insect will live between 5-7 days.

Cumulative Distribution Function

Properties of the CDF:

Assume that *f* is a probability density function, defined and continuous on an interval *[a,b]*. The left end *a* could be a real number or negative infinity; the right end *b* could be a real number or infinity. Denote by *F* the associated cumulative distribution function. Then

(1) $0 \le F(x) \le 1$ for all $x \in [a,b]$.

(2) $F(x)$ is continuous and non-decreasing.

(3) $\lim_{x \to a} F(x) = 0$ and $\lim_{x \to b} F(x) = 1$.

The Mean and the Variance

Definition:
Let *X* be a continuous random variable with probability density function *f(x)*, defined on an interval *[a,b]*.

The mean (or the expected value) of *X* is given by

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

The variance of *X* is

$$\mathrm{var}(X) = E\left[(X-\mu)^2\right] = \int_a^b (x-\mu)^2 f(x)dx$$

The Mean and the Variance

Example #24:

Consider the continuous random variable *X* given by the probability density function

$$f(x) = 0.3 + 0.2x \quad \text{for} \quad 0 \le x \le 2.$$

Find the probability that the values of *X* are at least one standard deviation above the mean.