Continuous probability distributions, chapter 5

Grethe Hystad
University of Arizona

March 24, 2011

Continuous random variables and their probability distributions, chapter 5.1

Definition

A random variable, X, is said to be continuous if there is a function, $f(x)$, called the probability density function, such that

- $f(x) \geq 0$ for all x
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$.

Continuous random variables and their probability distributions

- Notice that for a continuous random variable X,

$$P(X = a) = \int_a^a f(x)dx = 0 \text{ for any } a.$$

- Assign zero probability to any specific value.
- Hence,

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

- With discrete distributions, probability is associated with specific values.
- With continous distributions, positive probabilities are only associated with intervals.

Continuous random variables and their probability distributions

Example

The random variable X of the lifelengths of batteries is associated with a probability density function of the form,

$$f(x) = \begin{cases} \frac{1}{4} e^{-\frac{x}{4}}, & x > 0; \\ 0, & \text{elsewhere.} \end{cases}$$

with measurements in 100 hours.
(A) Find the probability that the life of a particular battery of this type is less than 100 or greater than 200 hours?
(B) Find the probability that a battery of this type lasts more than 500 hours given that it already has been in use for more than 300 hours?

Continuous random variables and their probability distributions

Solution

(A) The probability that the life of a particular battery of this type is less than 100 or greater than 200 hours is

$$\begin{aligned}
P((X<1)\cup(X\geq 2)) &= P(X<1)\cup P(X\geq 2)\\
&= \int_0^1 \tfrac{1}{4}e^{-\frac{x}{4}}\,dx + \int_2^\infty \tfrac{1}{4}e^{-\frac{x}{4}}\,dx\\
&= -e^{-\frac{x}{4}}|_0^1 - e^{-\frac{x}{4}}|_2^\infty\\
&= 1-e^{-\frac{1}{4}}+e^{-\frac{1}{2}} \approx 2.385.
\end{aligned}$$

Solution continue

Solution

(B) The probability that a battery of this type lasts more than 500 hours given that it already has been in use for more than 300 hours is

$$\begin{aligned} P(X > 5 \mid X > 3) &= \frac{P(X > 5)}{P(X > 3)} \\ &= \frac{\int_5^\infty \frac{1}{4} e^{-\frac{x}{4}}\, dx}{\int_3^\infty \frac{1}{4} e^{-\frac{x}{4}}\, dx} \\ &= \frac{e^{-\frac{5}{4}}}{e^{-\frac{3}{4}}} = e^{-\frac{1}{2}} = \approx 0.606. \end{aligned}$$

The distribution function

Definition

The distribution function for a continuous random variable X with the probability density function $f(x)$ is defined as

$$F(x) = P(X \leq x),$$

where

$$F(x) = \int_{-\infty}^{x} f(y)dy.$$

Notice that $F'(x) = f(x)$.

The distribution function

Example

Determine the distribution function for X, where its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{4}e^{-\frac{x}{4}}, & x > 0; \\ 0, & \text{elsewhere.} \end{cases}$$

The distribution function

Solution

$$
\begin{aligned}
F(x) &= P(X \leq x) \\
&= \int_0^x \tfrac{1}{4} e^{-\frac{y}{4}} \, dy \\
&= -e^{-\frac{y}{4}} \Big|_0^x \\
&= \begin{cases} 1 - e^{-\frac{x}{4}}, & x > 0; \\ 0, & \textit{elsewhere}. \end{cases}
\end{aligned}
$$

The distribution function

$F(x)$ is a distribution function for a continuous random variable iff

- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$
- $F(x)$ in nondecreasing, that is if $x < y$, then $F(x) \leq F(y)$.
- $F(x)$ is absolutely continuous over the whole real line.

The distribution function

Example

Let the random variable, X, be the time (in years) from a machine is serviced until it breaks down. Its distribution function is given as

$$F(x) = 1 - e^{-\frac{1}{2}x^{1.1}} \text{ for } x > 0.$$

(A) Find the probability that a randomly selected machine breaks down after at least 2 years.
(B) Find the probability density function of X.

The distribution function

Solution

(A) The probability that a randomly selected machine breaks down after at least 2 years is

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) \\ &= 1 - F(2) = e^{-\frac{1}{2}2^{1.1}} \approx 0.34 \end{aligned}$$

(B)

$$f(x) = F'(x) = \begin{cases} 0, & x < 0; \\ 0.55x^{0.1}e^{-\frac{1}{2}x^{1.1}}, & x \geq 0. \end{cases}$$

The distribution function

Example

Suppose a certain electronic system has a life length of X with a probability density function

$$f(x) = \begin{cases} cxe^{-\frac{x}{100}}, & x > 0; \\ 0, & \text{elsewhere.} \end{cases}$$

(A) Find the value of c that makes this function a valid probability density function.
(B) Find the cumulative distribution function for X.
(C) What is the probability that the system has a life lenght that exceeds 200 hours given that it exceeded 100 hours?

Solution

- *(A) Must solve the equation* $1 = \int_0^\infty cxe^{-\frac{x}{100}}dx$*. Solving the integral by integration by parts, we obtain* $c = \frac{1}{10000}$.

Solution

- *(A) Must solve the equation* $1 = \int_0^\infty cxe^{-\frac{x}{100}}dx$. *Solving the integral by integration by parts, we obtain* $c = \frac{1}{10000}$.
- *(B)*

$$F(x) = \frac{1}{(100)^2}\int_0^x ye^{-\frac{y}{100}}dy = \begin{cases} 1 - e^{-\frac{x}{100}}(\frac{x}{100} + 1), & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Solution

- *(A) Must solve the equation* $1 = \int_0^\infty cxe^{-\frac{x}{100}}dx$. *Solving the integral by integration by parts, we obtain* $c = \frac{1}{10000}$.
- *(B)*

$$F(x) = \frac{1}{(100)^2}\int_0^x ye^{-\frac{y}{100}}dy = \begin{cases} 1 - e^{-\frac{x}{100}}(\frac{x}{100}+1), & x \geq 0; \\ 0, & x < 0. \end{cases}$$

- *(C) The probability that the system has a life lenght that exceeds 200 hours given that it exceeded 100 hours is*

$$P(X > 200|X > 100) = \frac{P(X > 200)}{P(X > 100)} = \frac{1 - P(X \leq 200)}{1 - P(X \leq 100)}$$
$$= \frac{e^{-\frac{200}{100}}(\frac{200}{100}+1)}{e^{-\frac{100}{100}}(\frac{100}{100}+1)} = \frac{3e^{-2}}{2e^{-1}} = \frac{3}{2}e^{-1} \approx 0.552.$$

Expected values of continuous random variables, chapter 5.2

Definition

The expected value of a continuous random variable, X, that has a probability density function, $f(x)$, is given by

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx.$$

We assume the absolute convergence of all integrals so that the expected value exists.

Theorem

If X is a continuous random variable with probability distribution, $f(x)$, and if $g(x)$ is any real valued function of X, then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Definition

For a random variable X with probability density function, $f(x)$, the variance of X is given by

$$V(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = E(X^2) - \mu^2,$$

where $\mu = E(X)$.

Theorem

For constants a and b, we have

$$E(aX + b) = aE(X) + b$$

$$V(aX + b) = a^2 V(X).$$

Theorem

For any nonnegative continous random variable with distribution function $F(x)$ and finite mean $E(X)$,

$$E(X) = \int_0^\infty [1 - F(x)]dx$$

The proof of the last theorem is left as homework.

Expected values of continuous random variables, chapter 5.2

Example

Suppose the daily demand of a certain item, X, in a store sold by the pound and measured in hundreds of pounds, has a density function

$$f(x) = \begin{cases} \frac{x^2}{8}, & 0 \leq x \leq 2; \\ \frac{1}{3}, & 2 < x \leq 4; \\ 0, & \text{elsewhere} \end{cases}$$

Suppose the store's profit is \$30 for each 100 pounds sold (30 cents per pound if $X \leq 2$) and \$20 per 100 pounds if $X > 2$.
(A) Find the expected daily demand and variance of the daily demand.
(B) Find the store's expected profit for any given day.

Solution

(A)

$$\begin{aligned}\mu &= E(X)\\ &= \int_{-\infty}^{\infty} x f(x)\,dx\\ &= \int_0^2 x\frac{x^2}{8}\,dx + \int_2^4 x\frac{1}{3}\,dx\\ &= \tfrac{x^4}{32}\Big|_0^2 + \tfrac{x^2}{6}\Big|_2^4\\ &= 2.5\end{aligned}$$

solution continue

Solution

$$
\begin{aligned}
V(X) &= E(X^2) - \mu^2 \\
&= \int_0^2 x^2 \frac{x^2}{8} dx + \int_2^4 x^2 \frac{1}{3} dx - \left(\frac{5}{2}\right)^2 \\
&= \frac{139}{180}.
\end{aligned}
$$

Solution continue

(B) Let $g(X)$ denote the store's daily profit. Then

$$g(X) = \begin{cases} 3X, & 0 \le x \le 2; \\ 2X, & 2 < x \le 4 \end{cases}$$

Then the expected profit is

$$\begin{aligned} E(g(X)) &= \int_{-\infty}^{\infty} g(x) f(x) dx \\ &= \int_0^2 3x \frac{x^2}{8} + \int_2^4 2x \frac{1}{3} dx \\ &= 5.5. \end{aligned}$$

Thus, the expected daily profit of this item is 5.5 dollar.

Example

Example 5.15

The weekly repair cost, X, for a certain machine has a probability density function given by

$$f(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1; \\ 0, & \text{elsewhere} \end{cases}$$

with measurements in $\$100s$.

(A) Find the mean and variance of the distribution of repair costs.

(B) Find an interval within which these weekly repair costs should lie at least 75% of the time using Tchebysheff's Theorem.

(C) Find an interval within which these weekly repair costs lie exactly 75% of the time with exactly half of those not lying in the interval above the upper limit and the other half below the lower limit. Compare this interval to the one obtained in part (B).

Solution

(A)

$$\begin{aligned}\mu &= E(X)\\ &= \int_0^1 xf(x)dx\\ &= \int_0^1 x6x(1-x)dx\\ &= \int_0^1 (6x^2-6x^3)dx\\ &= \left(2x^3-\frac{3}{2}x^4\right)\Big|_0^1\\ &= \frac{1}{2}.\end{aligned}$$

Solution continue

Solution

$$\begin{aligned} V(X) &= E(X^2) - \mu^2 \\ &= \int_0^1 x^2 f(x)dx - \frac{1}{4} \\ &= \int_0^1 x^2 6x(1-x)dx - \frac{1}{4} \\ &= \int_0^1 (6x^3 - 6x^4)dx \\ &= \left(\frac{3}{2}x^4 - \frac{6}{5}x^5\right)\Big|_0^1 - \frac{1}{4} = \frac{1}{20}. \end{aligned}$$

Solution continue

(B) We have $\sigma = \sqrt{\frac{1}{20}}$. By Tchebysheff's inequality,

$$P\left(\left|X - \frac{1}{2}\right| < k\sqrt{\frac{1}{20}}\right) \geq 1 - \frac{1}{k^2} = 0.75.$$

Hence $k = 2$ Thus, these weekly repair costs lie in the interval

$$\left(\frac{1}{2} - 2\sqrt{\frac{1}{20}}, \frac{1}{2} + 2\sqrt{\frac{1}{20}}\right) = (0.053, 0.947)$$

at least 75% of the time.

Solution continue

(C) We want to find a and b such that $P(X < a) = 0.125$ and $P(X > b) = 0.125$. We have

$$0.125 = P(X < a) = \int_0^a 6x(1-x)dx$$

which gives $a = 0.22$. We have

$$0.125 = P(X > b) = 1 - P(X \leq b) = 1 - \int_0^b 6x(1-x)dx = 1 - 3b^2 - 2b^3$$

which gives $b = 0.78$.
So within the interval $(0.22, 0.78)$ the repair costs lie exactly 75% of the time with exactly half of those not lying in the interval above the upper limit and the other half below the lower limit.

The Uniform Distribution, chapter 5.3

Consider experiments that consists of observing

- events in a certain time interval such as phone calls coming into a call center
- particles that have a certain diameter
- distances of a point from the beginning of a line.
- Suppose that we know that one such event has occured in the interval (a, b).
- If X is the random variable for such experiments, we assume that X is equally likely to lie in any small subinterval within (a, b).
- X then has a uniform probability distribution.

The Uniform Distribution, chapter 5.3

Theorem

The uniform distribution:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & elsewhere \end{cases}$$

$$E(X) = \frac{a+b}{2} \text{ and } V(X) = \frac{(b-a)^2}{12}.$$

The Uniform Distribution, chapter 5.3

Theorem

The distribution function is:

$$F(x) = \begin{cases} 0 & x < a; \\ \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & x > b \end{cases}$$

If $(c, c+d) \subset (a, b)$, then

$$\begin{aligned} P(c \leq x \leq c+d) &= P(x \leq c+d) - P(x \leq c) \\ &= F(c+d) - F(c) = \frac{c+d-a}{b-a} - \frac{c-a}{b-a} = \frac{d}{b-a}. \end{aligned}$$

Thus, the probability depends only on the length d of the subinterval.

The expected value of The Uniform Distribution

The expected value of X is

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_a^b x\frac{1}{b-a}dx \\ &= \frac{1}{(b-a)}\frac{x^2}{2}|_a^b \\ &= \frac{b+a}{2} \end{aligned}$$

which is the midtpoint of the interval $[a, b]$.

The variance of The Uniform Distribution

The variance of X is

$$\begin{aligned}
V(X) &= E(X^2) - \mu^2 \\
&\int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 \\
&= \int_a^b x^2 \frac{1}{b-a} dx - \frac{(b+a)^2}{4} \\
&= \frac{1}{(b-a)} \left(\frac{b^3 - a^3}{3} \right) - \frac{(b+a)^2}{4} \\
&= \frac{b^2 + ab + a^2}{3} - \frac{(b+a)^2}{4} \\
&= \frac{1}{12}(b-a)^2.
\end{aligned}$$

The Poisson and the Uniform distribution

- Suppose the number of events that occur in an interval, (s, t), has a Poisson distribution.
- Suppose that exactly one of these events occured in the interval (a, b), where $a \geq s$ and $b \leq t$.
- Then the conditional probability distribution of the actual time of occurence for this event, given that is has occured, is uniform over (a, b).

Uniform probability density

Example

A customer is entering a store every 5 minutes. Let X be the time in minutes to the customer arrives. What is the distribution function, the expectation and the variance of X?

Solution

Solution

Uniform distribution:

$$f(x) = \begin{cases} \frac{1}{5} & 0 \le x \le 5; \\ 0, & otherwise \end{cases}$$

$$F(x) = \begin{cases} 0 & x < 0; \\ \int_0^x \frac{1}{5} dt = \frac{x}{5}, & 0 \le x \le 5; \\ 1, & x > 5 \end{cases}$$

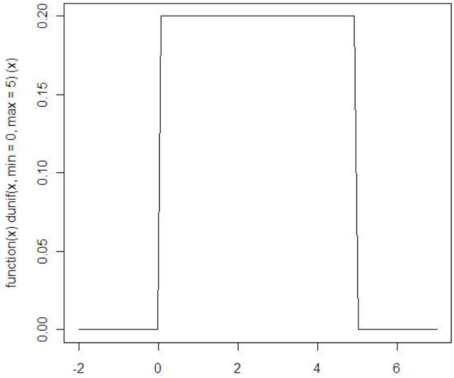
$a = 0,\ b = 5.$

$$E(X) = \frac{a+b}{2} = \frac{5}{2} \text{ and } V(X) = \frac{(b-a)^2}{12} = \frac{25}{12}.$$

Uniform probability density

In R:

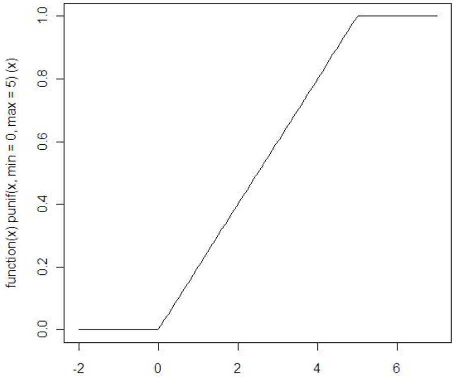
```
> plot(function(x) dunif(x,min=0,max=5),-2,7)
```



Uniform probability distribution

In R:

```
> plot(function(x) punif(x,min=0,max=5),-2,7)
```



The Uniform distribution

- **Example**

  Suppose that X has a uniform distribution on the interval $(0, a)$ for $a > 1$. Find $P(X > X^2)$.

The Uniform distribution

- **Example**

  Suppose that X has a uniform distribution on the interval $(0, a)$ for $a > 1$. Find $P(X > X^2)$.

- **Solution**

  $X > X^2$ *is equivalent to* $0 < X < 1$ *so*

  $$P(X > X^2) = P(0 < X < 1) = \int_0^1 \frac{1}{a} dx = \frac{1}{a}.$$