Continuous Probability Distributions, chapter 3.2

Grethe Hystad
University of Arizona

September 16, 2012

Continuous Random variables and their Probability Distributions, chapter 3.2

Definition

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

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

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 continuous distributions, positive probabilities are only associated with intervals.

Continuous random variables and their probability distributions

Example

The random variable X of the life-lengths 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 the life of a particular battery of this type is greater than 800 hours.
(B) Find the probability that the life of a particular battery of this type is less than 100 or greater than 200 hours

Continuous random variables and their probability distributions

Solution

(A) The probability that the life of a particular battery of this type is greater than 800 hours is

$$
\begin{aligned}
P(X > 8) &= \int_8^\infty \tfrac{1}{4} e^{-\frac{x}{4}}\, dx \\
&= -e^{-\frac{x}{4}} \Big|_8^\infty \\
&= e^{-2} \approx 0.135.
\end{aligned}
$$

Continuous random variables and their probability distributions

Solution

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

$$
\begin{aligned}
P((X<1)\cup(X\geq 2)) &= P(X<1)\cup P(X\geq 2)\\
&= \int_0^1 \tfrac{1}{4}e^{-\frac{x}{4}}\,dx + \int_2^\infty \tfrac{1}{4}e^{-\frac{x}{4}}\,dx\\
&= -e^{-\frac{x}{4}}\Big|_0^1 - e^{-\frac{x}{4}}\Big|_2^\infty\\
&= 1-e^{-\frac{1}{4}}+e^{-\frac{1}{2}} \approx 0.828.
\end{aligned}
$$

The distribution function

Definition

The (cumulative) distribution function (c.d.f.) of a continuous random variable X with the probability density function $f(x)$ is defined as

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

Notice that $F'(x) = f(x)$ (Fundamental Theorem of Calculus).

$F(x)$ cumulates all of the probability less than or equal to 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

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

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

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

(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 length 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 length 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

Definition

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

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

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

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

where $\mu = E(X)$.

Expected value as a linear operator

Theorem

- $\sigma^2 = Var(X) = E(X^2) - \mu^2$,
- *For constants a and b, we have* $E(aX + b) = aE(X) + b$
- $Var(aX + b) = a^2 Var(X)$.

Expected values of continuous random variables

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)

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

solution continue

Solution

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

Solution

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

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

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

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

with measurements in \$100s.

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

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

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

Percentile

Definition

The $(100p)^{th}$ percentile of a continuous distribution with p.d.f $f(x)$ is a number π_p such that $p = \int_{-\infty}^{\pi_p} f(x)\, dx = F(\pi_p)$.

- The $(50p)^{th}$ percentile, $m = \pi_{0.5}$, is called the **median**.
- The $(25p)^{th}$ percentile , $q_1 = \pi_{0.25}$, is called the **first quartile**.
- The $(75p)^{th}$ percentile , $q_3 = \pi_{0.75}$, is called the **third quartile**.

Percentile

Example

Let X have the p.d.f. $f(x) = 4x^3$ for $0 < x < 1$. Find the 40^{th} percentile, $\pi_{0.4}$.

Percentile

Example

Let X have the p.d.f. $f(x) = 4x^3$ for $0 < x < 1$. Find the 40^{th} percentile, $\pi_{0.4}$.

Solution

The distribution function of X is given by

$$F(x) = \int_0^x 4t^3 \, dt = x^4 \quad \textit{for} \quad 0 < x < 1.$$

The 40^{th} percentile, $\pi_{0.4}$, is given by $F(\pi_{0.4}) = 0.4$. We have $F(\pi_{0.4}) = \pi_{0.4}^4 = 0.4$ so $\pi_{0.4} \approx 0.795$. Thus $P(0 < X < 0.795) = \int_0^{0.795} f(x) \, dx = 0.4$.