Continuous Probability Functions II

Definitions

The *expected value* $E(X)$ of a random variable X with a continuous probability density function $f(x)$ on an interval $[a,b]$ is

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

The *variance* $\mathrm{Var}(X)$ of the same random variable is

$$\sigma^2 = \mathrm{Var}(X) = E((X-\mu)^2) = \int_a^b (x-\mu)^2 f(x)\, dx.$$

We could re-arrange this to get a slightly simpler formula:

$$\sigma^2 = \mathrm{Var}(X) = E(X^2) - \mu^2 = \int_a^b x^2 f(x)\, dx - \mu^2.$$

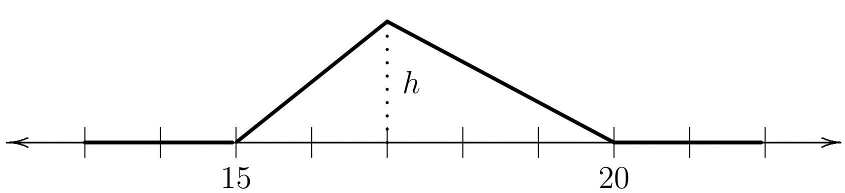
1 Consider the following probability density functions:

(a) $f(x) = \frac{1}{5}$ on $[0,5]$ (b) $f(x) = \frac{4}{3x^2}$ on $[1,4]$

For each of these functions, compute the following:

(i) $E(X)$ (ii) $\mathrm{Var}(X)$ (iii) $\Pr(X \le 3)$

2 Suppose a stock analyst believes that a stock's value in one week will vary between \$15 and \$20, with \$17 the most likely value. We can model this as a random variable X with a probability density function $f(x)$ shown by the following graph:



(a) Find the height h of this triangle. **Hint:** What's the total area of the triangle?

(b) Find the equation of $f(x)$.

(c) Use your answer to part (b) to find $E(x)$. Is it \$17? Should we choose a model so that $E(X) = \$17$?

(d) Find the variance $\sigma^2 = \mathrm{Var}(X)$ and the standard deviation σ.

Commonly Used Distributions

Distribution	Probability density function	$\mu = E(X)$	$\sigma^2 = \mathrm{Var}(X)$
General	$f(x)$ on $[a,b]$	$\int_a^b x f(x)\,dx$	$\int_a^b x^2 f(x)\,dx - \mu^2$
Uniform	$f(x) = \dfrac{1}{b-a}$ on $[a,b]$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Exponential	$f(x) = \lambda e^{-\lambda x}$ on $[0,\infty)$	${}^1/_\lambda$	${}^1/_{\lambda^2}$

3 Exponential models are often used for modelling waiting times. Usually we are given the mean $\mu = {}^1/_\lambda$, so $\lambda = {}^1/_\mu$. For example, suppose a manufacturer makes laptop batteries with an average charge life of 150 minutes (2.5 hours). Suppose we model the battery life with an exponential model X.

(a) What is λ?

(b) What is the probability that a battery charge will last less than 3 hours? More than 3 hours?

(c) What is $\Pr(X < 1.5 \text{ hours})$?

(d) What is $\Pr(X < \mu - \sigma)$?

4 Suppose a random number is selected at random from the interval $[0, 20]$. This determines a uniformly distributed random variable X.

(a) What is the mean μ and the standard deviation $\sigma = \sqrt{\mathrm{Var}(X)}$?

(b) Find the probability that X lies in the interval $[1, 5]$.

(c) What is $\Pr(X < \mu - \sigma)$? What is $\Pr(X < \mu - 2\sigma)$?

5 Often people will think that the mean μ has the property that $\Pr(X < \mu) = \Pr(X > \mu) = {}^1/_2$. (That is, they'll think that it's equally likely to be above or below the mean. Put another way, they'll confuse the mean with the median.) This is not always the case.

(a) Show, using the probability density function from Problem 3, that $\Pr(X < \mu) = \frac{e-1}{e} = 1 - \frac{1}{e} \approx 0.6321$.

(b) Show, using the probability density function from Problem 4, that $\Pr(X < \mu) = \frac{1}{2}$.

1 (a) (i) $E(X) = \frac{5}{2} = 2.5$ (ii) $\text{Var}(X) = \frac{25}{12} \approx 2.0833$ (iii) $\Pr(X \le 3) = \frac{3}{5} = 0.6$

(b) (i) $E(X) = \frac{4}{3}\ln(4) \approx 1.8484$ (ii) $\text{Var}(X) = 4 - \mu^2 \approx 0.5834$

(iii) $\Pr(X \le 3) = \frac{8}{9} \approx 0.8889$

2 (a) The area of the triangle is 1, so $h = \frac{2}{5} = 0.4$.

(b) If we assume that $f(x)$ has domain $[15, 20]$, then we need to find the equations of only two lines: one from $(x, y) = (15, 0)$ to $(x, y) = (17, 2/5)$ and the other from $(x, y) = (17, 2/5)$ to $(x, y) = (20, 0)$. Thus the equation for $f(x)$ is

$$f(x) = \begin{cases} \frac{1}{5}(x - 15) & \text{if } 15 \le x \le 17 \\ -\frac{2}{15}(x - 20) & \text{if } 17 < x \le 20. \end{cases}$$

(c) Using the general formula, we get $E(X) = \int_{15}^{20} x f(x)\, dx$. We split this up based on the definition of $f(x)$:

$$\begin{aligned} E(X) &= \int_{15}^{20} x f(x)\, dx = \int_{15}^{17} x f(x)\, dx + \int_{17}^{20} x f(x)\, dx \\ &= \int_{15}^{17} x \left(\frac{1}{5}(x - 15)\right) dx + \int_{17}^{20} x \left(-\frac{2}{15}(x - 20)\right) dx \\ &= \frac{1}{5}\int_{15}^{17} \left(x^2 - 15x\right) dx - \frac{2}{15}\int_{17}^{20} \left(x^2 - 20x\right) dx \\ &= \frac{1}{5} \cdot \frac{98}{3} - \frac{2}{15}(-81) = \frac{52}{3} = 17\,1/3 \approx 17.3333. \end{aligned}$$

Thus, as we might expect, our average is slightly to the right of the highest point on the curve.

(d) We'll use the formula $\sigma^2 = \text{Var}(X) = E(X^2) - \mu^2$. We begin by computing $E(X^2)$:

$$\begin{aligned} E(X^2) &= \int_{15}^{20} x^2 f(x)\, dx = \int_{15}^{17} x^2 f(x)\, dx + \int_{17}^{20} x^2 f(x)\, dx \\ &= \int_{15}^{17} x^2 \left(\frac{1}{5}(x - 15)\right) dx + \int_{17}^{20} x^2 \left(-\frac{2}{15}(x - 20)\right) dx \\ &= \frac{1}{5}\int_{15}^{17} \left(x^3 - 15x^2\right) dx - \frac{2}{15}\int_{17}^{20} \left(x^3 - 20x^2\right) dx \\ &= \frac{1}{5} \cdot 534 - \frac{2}{15}(-1460.25) = 301.5 \end{aligned}$$

The variance is then

$$\sigma^2 = \text{Var}(X) = E(X^2) - \mu^2 = 301.5 - \left(\frac{52}{3}\right)^2 = \frac{19}{18} \approx 1.0556.$$

The standard deviation is thus $\sigma = \sqrt{19/18} \approx 1.0274$.

3 (a) $\lambda = \frac{1}{2.5}$ or $\lambda = \frac{1}{150}$, depending whether we use hours or minutes.

(b) We will compute $\Pr(X < 3 \text{ hours})$:

$$\Pr(X < 3 \text{ hours}) = \int_0^3 \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^3 = -e^{-\frac{1}{2.5}\cdot 3} - \left(-e^0\right) = -e^{-1.2} + 1 \approx 0.6988.$$

The probability that a battery will last more than 3 hours is therefore $\Pr(X > 3 \text{ hours}) = 1 - \Pr(X < 3 \text{ hours}) = 1 - (-e^{-1.2} + 1) = e^{-1.2} \approx 0.3012$.

Of course, if we use $\lambda = \frac{1}{150}$, the computation would still look similar:

$$\Pr(X < 180 \text{ mins}) = \int_0^{180} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^{180} = -e^{-\frac{1}{150}\cdot 180} - \left(-e^0\right) = -e^{-1.2} + 1 \approx 0.6988.$$

(c) We will compute $\Pr(X < 1.5 \text{ hours})$:

$$\Pr(X < 1.5 \text{ hours}) = \int_0^{1.5} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^{1.5} = -e^{-\frac{1}{2.5}\cdot 1.5} - \left(-e^0\right) = -e^{-0.6} + 1 \approx 0.4512.$$

(d) Here $\mu = \frac{1}{\lambda} = \frac{1}{2.5}$ and $\sigma^2 = \frac{1}{2.5^2}$, so $\sigma = \frac{1}{2.5}$. Thus $\mu - \sigma = \frac{1}{2.5} - \frac{1}{2.5} = 0$ and

$$\Pr(X < \mu - \sigma) = \Pr(X < 0) = \int_0^0 \lambda e^{-\lambda x}\, dx = 0.$$

4 (a) $\mu = \frac{a+b}{2} = \frac{0+20}{2} = 10$.

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(20-0)^2}{12} = \frac{400}{12} = \frac{100}{3} \qquad \Longrightarrow \qquad \sigma = \sqrt{\frac{100}{3}} = \frac{10}{\sqrt{3}} \approx 5.7735.$$

(b) $\Pr(1 \le X \le 5) = \int_1^5 \frac{1}{20}\, dx = \frac{1}{20}x\Big|_1^5 = \frac{1}{20}\cdot 5 - \frac{1}{20}\cdot 1 = \frac{1}{5} = 0.2$.

(c) $\Pr(X < \mu - \sigma) = \Pr(X < 10 - \frac{10}{\sqrt{3}}) = \frac{10 - \frac{10}{\sqrt{3}}}{20} \approx 0.2113$.

$\Pr(X < \mu - 2\sigma) = 0$ because $\mu - 2\sigma = 10 - 2\cdot\frac{10}{\sqrt{3}} = 10 - \frac{20}{\sqrt{3}} \approx -1.5470$ is less than 0.

5 (a) This works for any exponential distribution (with mean $\mu = 1/\lambda$):

$$\Pr(X < \mu) = \int_0^{1/\lambda} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^{1/\lambda} = -e^{-\lambda\cdot\frac{1}{\lambda}} - \left(-e^0\right) = -e^{-1} + 1.$$

(b) Again, we compute for any uniform distribution:

$$\Pr(X < \mu) = \int_a^{(a+b)/2} \frac{1}{b-a}\, dx = \frac{1}{b-a}x\Big|_a^{(a+b)/2} = \frac{1}{b-a}\cdot\left(\frac{a+b}{2} - a\right) = \frac{1}{b-a}\cdot\frac{b-a}{2} = \frac{1}{2}.$$

This messing around with a's and b's confuses the issue. It's easier with numbers (I'll use Problem 4 numbers):

$$\Pr(X < \mu) = \int_0^{10} \frac{1}{20}\, dx = \frac{1}{20}x\Big|_0^{10} = \frac{1}{20}\cdot(10 - 0) = \frac{1}{2}.$$