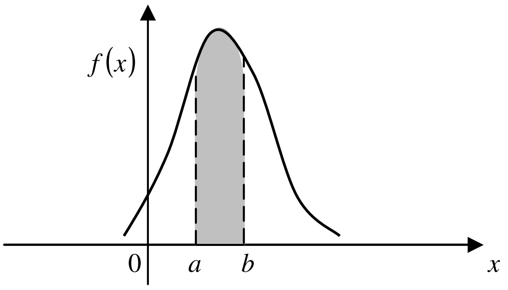
Free maths self-help revision

Continuous random variables and probability density functions

In many experiments a random variable can take any real value. The exact value of a random variable in any attempt of an experiment can never be determined due to the limitations of the measuring instruments and the person taking the measurement, e.g. measuring the height H of a randomly chosen person from a large number of people. It is more appropriate to specify a range of values (i.e. an interval, e.g. 170.0±0.5 cm) for the height rather than a single value.
Hence the assigned probability for $H = h$ *cm is zero, i.e.* $\Pr(H = h) = 0$, *and* $0 \le \Pr(a < H < b) \le 1$.

A random variable that fits the above descriptions is called a ***continuous*** *random variable. The* ***probability distribution*** *of a continuous random variable is given by its* ***probability density function*** $f(x)$ *such that* $\Pr(a < X < b) = \int_a^b f(x)dx$, *i.e. the area under the curve* $f(x)$ *in the interval* (a, b).



Notes: (1) For $f(x)$ *to be a probability density function of random variable X, it must satisfy the conditions:* $f(x) \ge 0$ *for all real x, and* $\int_{-\infty}^{\infty} f(x)dx = 1$.
(2) Do not confuse probability density function $f(x)$ *with probability function* $p(x)$ *discussed in previous lessons.* $f(x)$ *is for continuous random variable and the area under the graph of* $f(x)$ *gives the probability.* $p(x)$ *is for discrete random variable and direct substitution of a value of x gives the probability.*

Example 1 Random variable *X* has a probability density function given by $f(x) = \begin{cases} 1 - |x - 1| & for \quad 0 \le x \le 2 \\ 0 & elsewhere \end{cases}$.

Find (a) $\Pr(x = 1)$ (b) $\Pr(x < 1)$ (c) $\Pr(0.5 < x < 1.5)$
(d) $\Pr(x \ge 0.5 \mid x < 1.5)$.

(a) $\Pr(x = 1) = 0$

(b) $\Pr(x < 1) = \int_{-\infty}^{1} (1 - |x - 1|)dx = \int_0^1 (1 - |x - 1|)dx = 0.5$

(c) $\Pr(0.5 < x < 1.5) = \int_{0.5}^{1.5} (1 - |x - 1|)dx = 0.75$

(d) $\Pr(x \ge 0.5 \mid x < 1.5) = \dfrac{\Pr(x \ge 0.5 \cap x < 1.5)}{\Pr(x < 1.5)} = \dfrac{0.75}{0.875} = \dfrac{6}{7}$

Example 2 Random variable *X* has a distribution given by the probability density function

$f(x) = \begin{cases} ax(5 - x) & if \quad 0 \le x \le 5 \\ 0 & elsewhere \end{cases}$ where $a > 0$. Find *a*.

$\int_0^5 ax(5 - x)dx = a\int_0^5 (5x - x^2)dx = 1$, $a\left[\dfrac{5x^2}{2} - \dfrac{x^3}{3}\right]_0^5 = 1$, $a = \dfrac{6}{125}$

Example 3 Find *a*, *b* and *k* such that $f(a) = f(b) = 0$ and

$f(x) = \begin{cases} -k(x^2 - 2x) & for \quad a \le x \le b \\ 0 & elsewhere \end{cases}$

is a probability density function, where $k > 0$.

$f(x) = -k(x^2 - 2x) = 0$ $\therefore$ $x^2 - 2x = x(x - 2) = 0$, $x = 0, 2$
$\therefore$ $a = 0$, $b = 2$

$\int_0^2 -k(x^2 - 2x)dx = 1$, $\left[-k\left(\dfrac{x^3}{3} - x^2\right)\right]_0^2 = 1$, $-k\left(\dfrac{8}{3} - 4\right) = 1$ $\therefore$ $k = \dfrac{3}{4}$

Example 4 Find α such that $f(x) = \begin{cases} \alpha - e^{-x} & for \quad 0 \le x \le 1 \\ 0 & elsewhere \end{cases}$ is a probability density function. Hence find $\Pr(0.5 < x < 1)$.

$\int_0^1 (\alpha - e^{-x})dx = [\alpha x + e^{-x}]_0^1 = \alpha + e^{-1} - 1 = 1$ $\therefore$ $\alpha = 2 - e^{-1}$

$\Pr(0.5 < x < 1) = \int_{0.5}^{1} (2 - e^{-1} - e^{-x})dx = [(2 - e^{-1})x + e^{-x}]_{0.5}^1$
$= 2 - e^{-1} + e^{-1} - 0.5(2 - e^{-1}) - e^{-0.5} = 1 + 0.5e^{-1} - e^{-0.5}$

Calculation of expected value, variance and standard deviation of a continuous random variable

For continuous random variable X with probability density function $f(x)$,

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

$Var(X) = E((X - \mu_X)^2) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)dx$ *or*

$Var(X) = E(X^2) - (\mu_X)^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - (\mu_X)^2$,

$\sigma_X = \sqrt{Var(X)}$.

For many continuous probability distributions,
$\Pr(\mu - \sigma < X < \mu + \sigma) \approx 0.68$, *and*
$\Pr(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95$.

Calculation of median and mode of a continuous random variable

The median of continuous random variable X with probability density function $f(x)$ *is defined as the value m of X such that* $\int_{-\infty}^{m} f(x)dx = 0.5$. *Hence* $\int_m^{\infty} f(x)dx = 0.5$.

The mode of continuous random variable X with probability density function $f(x)$ *is the value M of X such that* $f(M) \ge f(x)$ *for all real x. The equal sign suggests that there may be more than one mode in the distribution. If* $f(x)$ *is differentiable for all x, then it is possible in some cases to find the mode by differentiation (maximum of* $f(x)$*).*

If the probability density function is symmetrical about a mode M, then the median m, the mean μ *and M have the same value.*

Example 5 Consider the probability density function $f(x)=\frac{2}{\pi}\sin^2 x$ for $0\le x\le \pi$ and $f(x)=0$ elsewhere.

(a) Show that $f(x)$ satisfies the requirements of a probability density function.
(b) Find $\Pr(1<x<2)$.
(c) Find $\Pr(x<2 \mid x>1)$.
(d) Find the mean, median and mode of *X*.
(e) Find the variance of *X*.
(f) Verify that $\Pr(\mu-2\sigma<x<\mu+2\sigma)\approx 0.95$.

(a) Since $\sin^2 x\ge 0$ for all $x\in[0,\pi]$, $\therefore f(x)\ge 0$ for all $x\in[0,\pi]$. Elsewhere $f(x)=0$. $\therefore f(x)\ge 0$ for all real *x*.

Use CAS to obtain $\int_0^{\pi}\frac{2}{\pi}\sin^2 x dx=1$. $\therefore \int_{-\infty}^{\infty} f(x)dx=1$.

(b) $\Pr(1<x<2)=\int_1^2 \frac{2}{\pi}\sin^2 x dx=0.5835$ by CAS

(c) $\Pr(x<2 \mid x>1)=\frac{\Pr(x<2\cap x>1)}{\Pr(x>1)}=\frac{\Pr(1<x<2)}{\Pr(x>1)}\approx 0.7061$

by CAS.

(d) Since $f(x)=\frac{2}{\pi}\sin^2 x$, $0\le x\le\pi$ is symmetrical about $\frac{\pi}{2}$,

$\therefore$ the mean = median = mode = $\frac{\pi}{2}$.

(e) $Var(X)=E(X^2)-(\mu_X)^2=\int_{-\infty}^{\infty}x^2 f(x)dx-(\mu_X)^2$

$$Var(X)=\int_0^{\pi}x^2\left(\frac{\pi}{2}\sin^2 x\right)dx-\left(\frac{\pi}{2}\right)^2=2.7899-2.4674=0.3225$$

and $\sigma_X=\sqrt{0.3225}=0.5679$.

(f) $\mu-2\sigma=\frac{\pi}{2}-2\times 0.5679=0.4350$,

$\mu+2\sigma=\frac{\pi}{2}+2\times 0.5679=2.7066$,

$$\Pr(0.4350<x<2.7066)\approx\int_{0.4350}^{2.7066}\frac{2}{\pi}\sin^2 x dx\approx 0.9664$$

Example 6 Consider the probability density function $f(x)=30x^4(1-x)$ for $0\le x\le 1$ and $f(x)=0$ elsewhere.

(a) Show that $f(x)$ satisfies the requirements of a probability density function. (b) Find $\Pr(x<0.5)$. (c) Find the mean of *X*. (d) Find the variance of *X*. (e) Use calculus to find the median of *X*. (f) Find the mode of *X*.

(a) For $0\le x\le 1$, $x^4\ge 0$ and $(1-x)\ge 0$, $\therefore f(x)\ge 0$. Elsewhere, $f(x)=0$. $\therefore f(x)\ge 0$ for all real *x*.

Use CAS to obtain $\int_0^1 30x^4(1-x)dx=1$. $\therefore \int_{-\infty}^{\infty} f(x)dx=1$.

By calculus $\int_0^1 30x^4(1-x)dx=\int_0^1 30(x^4-x^5)dx$

$$=\left[30\left(\frac{x^5}{5}-\frac{x^6}{6}\right)\right]_0^1=30\times\frac{1}{30}=1.$$

(b) By CAS or calculus

$$\Pr(x<0.5)=\int_0^{0.5}30x^4(1-x)dx=\int_0^{0.5}30(x^4-x^5)dx=0.1094$$

(c) By CAS/calculus $E(X)=\int_0^1 x30(x^4-x^5)dx=\int_0^1 30(x^5-x^6)dx$

$$=\left[30\left(\frac{x^6}{6}-\frac{x^7}{7}\right)\right]_0^1=30\times\frac{1}{42}=0.7143$$

(d) $Var(X)=\int_0^1 x^2 30(x^4-x^5)dx-0.7143^2$

$$=\int_0^1 30(x^6-x^7)dx-0.7143^2$$

$$=\left[30\left(\frac{x^7}{7}-\frac{x^8}{8}\right)\right]_0^1-0.7143^2=30\times\frac{1}{56}-0.7143^2=0.0255.$$

(e) Let $0<m<1$ be the median of *X*. $\therefore \Pr(x<m)=0.5$,

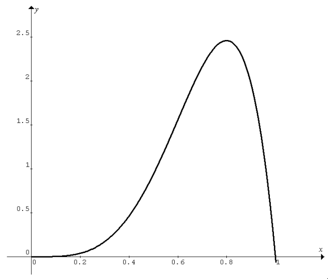
$\int_0^m 30(x^4-x^5)dx=0.5$.

$$\therefore\left[30\left(\frac{x^5}{5}-\frac{x^6}{6}\right)\right]_0^m=30\left(\frac{m^5}{5}-\frac{m^6}{6}\right)=0.5,\ \therefore 12m^5-10m^6=1.$$

Solve by CAS $m=0.7356$.

(f) $f(x)=30x^4(1-x)$ for $0\le x\le 1$

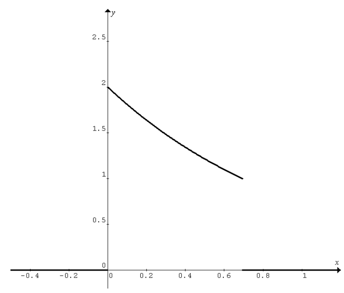
Sketch $f(x)$. This step is necessary to find out whether the probability density function has an endpoint maximum or a local maximum.



To find the maximum of $f(x)$, let $f'(x)=30(4x^3-5x^4)=0$, $\therefore x^3(4-5x)=0$, $\therefore x=0.8$, hence the mode is 0.8.

Example 7 Find the mode of *X* with probability density function

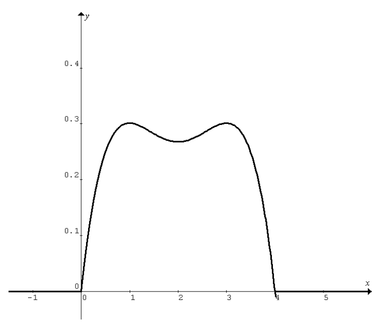
$$f(x)=\begin{cases}0 & if \quad x<0\\ 2e^{-x} & if \quad 0\le x\le \log_e 2\\ 0 & if \quad x>\log_e 2\end{cases}$$



$f(x)=2e^{-x}$ for $0\le x\le\log_e 2$ has a left endpoint maximum, hence the mode of *X* is $x=0$.

Example 8 Find the mode of *X* with probability density function

$$f(x)=\begin{cases}0 & if \quad x<0\\ \frac{15}{112}\left(6x-\frac{11}{2}x^2+2x^3-\frac{1}{4}x^4\right) & if \quad 0\le x\le 4\\ 0 & if \quad x>4\end{cases}$$



X has two modes, $x=1$ and $x=3$, by CAS.
Both values satisfy the requirement $f(M)\ge f(x)$ *for all real x.*

Questions: Next page

1. $f(x)=\begin{cases}\dfrac{x}{24} & 4\le x\le 8\\ 0 & elsewhere\end{cases}$ is the probability density function of a continuous random variable *X*.

(a) Find $\Pr(x<6)$. (b) Find the value of *a* if $\Pr(x\ge a)=\dfrac{5}{16}$.

2. $f(x)=\begin{cases}k\cos(\pi x) & -0.5\le x\le 0.5\\ 0 & elsewhere\end{cases}$ is the probability density function of a continuous random variable *X*.

(a) Find the value of *k*. (b) Find $\Pr(x\le 0 \mid x\le 0.25)$.

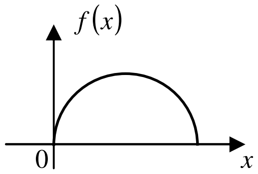
3. $f(x)=\begin{cases}0.5\times|1-0.5x| & 0\le x\le 4\\ 0 & elsewhere\end{cases}$ is the probability density function of a continuous random variable *X*.

(a) Find $\Pr(0.5<x<3.5)$. (b) Find $\Pr(x\ge 0.5 \mid x\le 3.5)$.

4. $f(x)=\begin{cases}a\sin(2x) & 0\le x\le \dfrac{\pi}{2}\\ 0 & elsewhere\end{cases}$ is the probability density function of a continuous random variable *X*. (a) Find the value of *a*. (b) Find the value of *b* such that $\Pr(x\le b)=0.25$.

5. $f(x)=\begin{cases}\dfrac{k}{x} & 1\le x\le e^2\\ 0 & elsewhere\end{cases}$ is the probability density function of a continuous random variable *X*. (a) Find the value of *k*.
(b) Find the mean of *X*. (c) Find the mode of *X*.

6. The graph of the probability density function of a continuous random variable *X* is a semicircle as shown below.
(a) Find *a* such that $\Pr(x\ge a)=0.5$
(b) Find the mean, median and the mode of *X*.



7. Explain why $f(x)=\begin{cases}\dfrac{4}{3}\left(1-\dfrac{1}{4}x^2\right) & 0\le x\le 3\\ 0 & elsewhere\end{cases}$ cannot be a probability density function.

8. Find the mean and standard deviation of *X* with probability density function $f(x)=\begin{cases}ax(x-1) & 0\le x\le 1\\ 0 & elsewhere\end{cases}$. Hence evaluate $\Pr(\mu-1.5\sigma\le x\le \mu+1.5\sigma)$.

9. Find the mode and median of *X* with probability density function

$$f(x)=\begin{cases}\dfrac{225}{1756}\left(\dfrac{x^5}{3}-\dfrac{7x^4}{12}-\dfrac{11x^3}{6}+\dfrac{317x^2}{150}+2x+1\right) & -2\le x\le 2\\ 0 & elsewhere\end{cases}.$$

10. The newspaper will be delivered to you between 6:30 and 7:30 am every day. And you will leave for school between 7:00 and 8:00 every morning. What's the probability that you can get the newspaper before you leave?

Numerical, algebraic and worded answers.

5. (a) 0.5 (b) $0.5(e^2-1)$ (c) 1 8. $\mu=0.5$, $\sigma=1/\sqrt{20}$, $\Pr(\mu-1.5\sigma\le x\le\mu+1.5\sigma)=0.8553$

9. mode = –1.5222, median = 0.1845 6. $a=\sqrt{2/\pi}$, mean = median = mode = $\sqrt{2/\pi}$ 3. (a) 0.5625 (b) 0.72

10. 0.875 7. $f(x)<0$ when $2<x\le 3$ 1. (a) 5/12 (b) $a=7$ 2. (a) $k=\pi/2$ (b) $\sqrt{2}/\left(1+\sqrt{2}\right)$ 4. (a) $a=1$ (b) $\pi/6$