Continuous Random Variables I Cheat Sheet

Describing Continuous Random Variables

There are some quantities which cannot be described accurately by a discrete random variable; for example heights, weights, or times. To handle this, continuous random variables can be used. Many properties of discrete random variables are shared with continuous random variables. However, there are also some differences. For discrete random variables, a probability can be associated to obtaining an exact value. Continuous random variables take values within an interval; so instead, the probability of the continuous random variable lying within a specified interval is considered. The probabilities of discrete random variables can be found using a probability mass function (pmf). Similarly, the probabilities of continuous random variables can be found using a probability density function (PDF). The probability of a continuous random variable X with PDF $f(x)$ lying in the interval $a < X < b$ is given by integrating the PDF between the two values:

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

Notice that $P(a < X < b) = P(a \leq X \leq b)$ for all continuous random variables in all intervals since the probability of X being exactly a or b is zero, this can be thought of as the "area" of a vertical line being exactly zero.

There are two conditions for $f(x)$ to be a probability density function:

- $\int_{-\infty}^{\infty} f(x)\,dx = 1$
- $f(x) \geq 0$ for all x

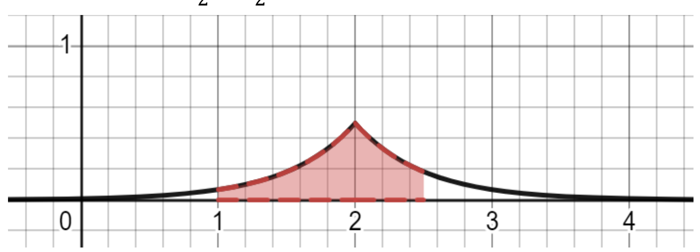
In words, this means the total probability over all cases must sum to 1 and that no probability can ever be negative. Notice that the limits of $-\infty$ and ∞ express the fact that the continuous random variable can take any real value, but in reality, they can be replaced with the lowest and highest value the continuous random variable takes. Also observe that as long as the probability density functions still satisfy the two conditions, they can be defined piecewise.

Example 1: Consider the probability density functions below.

a) $f(x) = \begin{cases} kx^2 & 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$ b) $f(x) = \begin{cases} \frac{k}{x^3} & 1 < x < 5 \\ 0 & \text{otherwise} \end{cases}$ c) $f(x) = \begin{cases} ke^{2(x-2)} & x \leq 2 \\ ke^{-2(x-2)} & x > 2 \end{cases}$

i) Solve for k.

ii) Find $P(1 < x < 2.5)$

a) i) Integrate pdf and set equal to 1 and solve for k.	$\int_1^3 kx^2\,dx = 1 \Rightarrow \left[\frac{kx^3}{3}\right]_1^3 = 1 \Rightarrow 9k - \frac{k}{3} = 1 \Rightarrow k = \frac{3}{26}$
a) ii) Integrate the pdf between $x = 1$ and $x = 2.5$.	$P(1 < x < 2.5) = \int_1^{2.5} \frac{3}{26}x^2\,dx = 3\left[\frac{x^3}{26}\right]_1^{2.5} = \frac{9}{16} = 0.563\ (3\text{ s.f.})$
b) i) Integrate pdf and set equal to 1 and solve for k.	$\int_1^5 \frac{k}{x^3}\,dx = 1 \Rightarrow \left[-\frac{k}{2x^2}\right]_1^5 = 1 \Rightarrow \frac{-k}{50} + \frac{k}{2} = 1 \Rightarrow k = \frac{25}{12}$
b) ii) Integrate the pdf between $x = 1$ and $x = 2.5$.	$P(1 < x < 2.5) = \int_1^{2.5} \frac{25}{12x^3}\,dx = \left[-\frac{25}{24x^2}\right]_1^{2.5} = \frac{7}{8} = 0.875\ (3\text{ s.f.})$
c) i) Integrate each part of pdf separately and set equal to 1 and solve for k.	$\int_{-\infty}^2 ke^{2(x-2)}\,dx + \int_2^{\infty} ke^{-2(x-2)}\,dx = 1 \Rightarrow \left[\frac{ke^{2(x-2)}}{2}\right]_{-\infty}^2 + \left[-\frac{ke^{-2(x-2)}}{2}\right]_2^{\infty} = 1$ $\Rightarrow \left(\frac{k}{2} - 0\right) + \left(0 - \left(-\frac{k}{2}\right)\right) = 1 \Rightarrow \mathrm{k} = 1$
c)ii) Integrate the first piece of the pdf between $x = 1$ and $x = 2$ and add this to the second piece integrated between $x = 2$ and $x = 2.5$.	$P(1 < x < 2.5) = \int_1^2 e^{2(x-2)}\,dx + \int_2^{2.5} e^{-2(x-2)}\,dx$ $= \left[\frac{e^{2(x-2)}}{2}\right]_1^2 + \left[-\frac{e^{-2(x-2)}}{2}\right]_2^{2.5}$ $= \left(\frac{1}{2} - \frac{e^{-2}}{2}\right) + \left(-\frac{e^{-1}}{2} - \left(-\frac{1}{2}\right)\right)$ $= 1 - \frac{e^{-2}}{2} - \frac{e^{-1}}{2} = 0.748\ (3\text{s.f.})$ 

Distributions of Random Variables That Are Part Discrete and Part Continuous

Some random variables may take a combination of discrete and continuous values. For example, the waiting time for a bus to come may be measured precisely for the first 15 minutes (making this part continuous) and then in 5 minutes intervals after this (making this part discrete). These variables can be handled by separating the different parts and using the relevant methods on each part. So, for calculating a probability, integrate between the two endpoints for the continuous part and sum between the two endpoints for the discrete part. This is demonstrated in the example below.

Example 2: Consider the mixed random variable X. It is continuous between $x = 0$ and $x = 3$, and takes the discrete values $x = 4,5,6,7,8$.
Its distribution is defined as: $f(x) = \frac{1}{27}x^3$ for $0 \leq x \leq 3$ and $P(X = x) = \frac{1}{20}$ for $x = 4,5,6,7,8$.
Find $P(2.5 < x < 6)$.

Split into continuous and discrete parts. Observe for the discrete part the only bits with non-zero probability are $x = 4$ and $x = 5$.	$P(2.5 < x < 6) = P(2.5 < x < 3) + P(3 < x < 6)$ $= P(2.5 < x < 3) + P(x = 4) + P(x = 5)$
Calculate the continuous part by integrating.	$P(2.5 < x < 3) = \int_{2.5}^3 \frac{1}{27}x^3 dx = \left[\frac{1}{108}x^4\right]_{2.5}^3 = \frac{3^4}{108} - \frac{(2.5)^4}{108} = \frac{671}{1728}$
Calculate the discrete part by summing.	$P(x = 4) + P(x = 5) = \frac{1}{20} + \frac{1}{20} = \frac{1}{10}$
Add the continuous and discrete parts to obtain the total probability.	$P(2.5 < x < 6) = \frac{671}{1728} + \frac{1}{10} = \frac{4219}{8640} = 0.488\ (3\text{ s.f.})$

Median and Quartiles for a Given Probability Density Function

The median m of a random variable X is the point at which there is an equal probability of X lying on either side of m. If X has the probability density function $f(x)$ then m satisfies:

$$\int_{-\infty}^m f(x)\,dx = \frac{1}{2}$$

Quartiles Q_1 and Q_3 can be defined similarly, so:

$$\int_{-\infty}^{Q_1} f(x)\,dx = \frac{1}{4}, \quad \int_{-\infty}^{Q_3} f(x)\,dx = \frac{3}{4}$$

Example 3: Consider the following probability density function:

$$f(x) = \begin{cases} \frac{3}{50}(x^2 + 2x + 5) & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the median m of $f(x)$.

b) Find the lower and upper quartiles Q_1 and Q_3 of $f(x)$.

a) Substitute $f(x)$ into the formula for the median with the lower limit as 0 since the pdf is 0 for all values less than 0.	$\int_0^m \frac{3}{50}(x^2 + 2x + 5)\,dx = \frac{1}{2}$
Integrate and substitute in limits forming a cubic.	$\frac{1}{50}m^3 + \frac{3}{50}m^2 + \frac{3}{10}m - \frac{1}{2} = 0$
Solve the cubic using a calculator to obtain m, rejecting complex roots.	$\Rightarrow m = 1.24\ (3\text{ s.f.})$
b) Substitute $f(x)$ into the formula for the lower & upper quartiles.	$\int_0^{Q_1} \frac{3}{50}(x^2 + 2x + 5)\,dx = \frac{1}{4}$ & $\int_0^{Q_3} \frac{3}{50}(x^2 + 2x + 5)\,dx = \frac{3}{4}$
Integrate and substitute in limits, forming a cubic equation.	$\frac{1}{50}{Q_1}^3 + \frac{3}{50}{Q_1}^2 + \frac{3}{10}Q_1 - \frac{1}{4} = 0$ & $\frac{1}{50}{Q_3}^3 + \frac{3}{50}{Q_3}^2 + \frac{3}{10}Q_3 - \frac{3}{4} = 0$
Solve the cubic using a calculator to obtain m, rejecting complex roots.	$\Rightarrow Q_1 = 0.709\ (3\text{ s.f.})$ & $Q_3 = 1.65\ (3\text{ s.f.})$
Sketch a graph of the probability density function with the quartiles and median to check that the answers look reasonable.