Continuous Random Variables III Cheat Sheet

AQA A Level Further Maths: Statistics

Cumulative Distribution Function (A Level Only)

A cumulative distribution function, (CDF), measures the probability of the variable taking a value which is less than or equal to a given value. For a cumulative distribution function $F(x)$,

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

Relations Between the Probability Density and Cumulative Distribution Functions

For a continuous random variable X with the probability density function (pdf) $f(x)$, its cumulative distribution function can be written as $F(x)$ and can be found by integrating $f(x)$ between $-\infty$ and x. The cumulative probability of X taking the values within the range $P(a \leq x \leq b)$ can be found by using a and b as the limits in the integration. When a and b are the smallest and largest values which X can take, $F(x)$ should equal to 1 as probability always adds up to 1.

$$F(x) = P(a \leq x \leq b) = \int_a^b f(x)\,dx$$

This is the same formula used to find the median, lower quartile and upper quartile. It can also be used for calculating percentiles, by setting the value of $F(x)$ to the desired percentile and finding the value of b.

The pdf can be found from a given cdf by differentiation:

$$f(x) = \frac{d}{dx}F(x).$$

Example 1: The continuous random variable X has the probability density function:

$$f(x) = \begin{cases} \frac{3}{50}(x^2 - 4x + 5) & 0 < x \leq k \\ 0 & \text{otherwise} \end{cases}$$

Find the cumulative distribution function of X and the value of k.

Integrate $f(x)$, using k and 0 as the limits.	$\begin{aligned} F(x) &= \int_0^k \frac{3}{50}(x^2 - 4x + 5)\,dx = \frac{3}{50}\left[\frac{x^3}{3} - \frac{4x^2}{2} + 5x\right]_0^k \\ &= \frac{3}{50}\left(\frac{k^3}{3} - 2k^2 + 5k\right) \end{aligned}$
Equate $F(x)$ to 1 and solve for k.	$\begin{aligned} \frac{3}{50}\left(\frac{k^3}{3} - 2k^2 + 5k\right) &= 1 \\ k^3 - 6k^2 + 15k &= 50 \\ k &= 5 \end{aligned}$
The cumulative distribution function can be written as:	$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{3}{50}\left(\frac{x^3}{3} - 2x^2 + 5k\right) & 0 < x \leq 5 \\ 1 & x \geq 5 \end{cases}$

Example 2: The continuous random variable X has the cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{4} & 0 \leq x < 1 \\ \frac{x}{5} + \frac{1}{20} & 1 \leq x < \frac{19}{4} \\ 1 & x \geq \frac{19}{4} \end{cases}$$

Find the probability density distribution and the median of X.

Find the pdf by differentiating each part of the cdf.	$f(x) = \begin{cases} \frac{1}{4} & 0 \leq x < 1 \\ \frac{1}{5} & 1 \leq x < \frac{19}{4} \\ 0 & \text{otherwise} \end{cases}$
To find the median, find the value of x when $F(x) = 0.5$. Note that the maximum value of $F(x)$ for $0 \leq x < 1$ is $\frac{1}{4}$, so the median must lie within $1 \leq x < \frac{19}{4}$.	$\begin{aligned} \frac{x}{5} + \frac{1}{20} &= 0.5 \\ x &= \frac{9}{4} \end{aligned}$

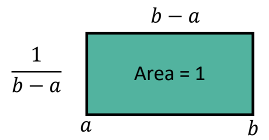
The Rectangular Distribution (A Level Only)

Applicability of Modelling with the Rectangular Distribution

A continuous random variable X is said to have a rectangular distribution if it has uniform distribution. This means there is equal probability of X taking a value within any ranges of the same width.

Calculating Probabilities with a Rectangular Distribution

The graph below represents the probability distribution of variable X which follows a rectangular distribution over $[a, b]$, where $b > a$.



Since the total area under graph is 1 and the base length of the rectangular region is $b - a$, the probability can be calculated by the following:

$$f(x) = \frac{1}{b-a}.$$

The cumulative distribution function, derived using the area of the rectangle, is given by:

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

Proof of Mean, Variance and Standard Deviation of the Rectangular Distribution

The mean and variance of X, which is a random variable following a rectangular distribution over $[a, b]$, is given by:

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

Example 3: For the random continuous variable X following a rectangular distribution $[a, b]$, a) show that $E(X) = \frac{a+b}{2}$ and b) $Var(X) = \frac{(b-a)^2}{12}$. Given that $a = 1$ and $b = 5$, find c) the standard deviation of X and d) $P(3.5 < x \leq 4.7)$.

a) Find $E(X)$ using $\int_{-\infty}^{\infty} x\, f(x)\,dx$ and $f(x) = \frac{1}{b-a}$. Use a and b as the limits for integration.	$\begin{aligned} E(X) &= \int_a^b x \times \frac{1}{b-a}dx \\ &= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b \\ &= \frac{1}{b-a} \times \frac{b^2 - a^2}{2} \\ &= \frac{a+b}{2} \end{aligned}$
b) Find $E(X^2)$ using $\int_{-\infty}^{\infty} x^2 f(x)\,dx$ and $f(x) = \frac{1}{b-a}$. Use a and b as the limits for integration.	$\begin{aligned} E(X^2) &= \int_a^b x^2 \times \frac{1}{b-a}dx \\ &= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b \\ &= \frac{1}{b-a} \times \frac{b^3 - a^3}{3} \\ &= \frac{a^2 + ab + b^2}{3} \end{aligned}$
Find $Var(X)$ using $Var(X) = E(X^2) - (E(X))^2$.	$\begin{aligned} Var(X) &= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{a^2 + ab + b^2}{3} - \left(\frac{a^2 + 2ab + b^2}{4}\right) \\ &= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} \\ &= \frac{a^2 - 2ab + b^2}{12} \\ &= \frac{(b-a)^2}{12} \end{aligned}$
c) Calculate the standard deviation from $Var(X)$, substituting the values of a and b.	$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(5-1)^2}{12}} = 1.15 \text{ (3 s.f.)}$
d) Find $P(3.5 < x \leq 4.7)$.	$\begin{aligned} P(3.5 < x \leq 4.7) &= F(4.7) - F(3.5) \\ &= \frac{4.7-1}{5-1} - \frac{3.5-1}{5-1} \\ &= \frac{4.7-3.5}{4} \\ &= 0.3 \end{aligned}$

Expectation and Variance of the Sum of Two Independent Random Variables (Discrete or Continuous)

When the independent random variables X and Y are combined, the expectation and variance are given by:

$$E(X+Y) = E(X) + E(Y)$$

$$Var(X+Y) = Var(X) + Var(Y)$$

Notice that this holds even if one is continuous and the other is discrete.

Example 4: The random variable X has the probability density function:

$$f(x) = \begin{cases} 2k & 1 < x < 3 \\ k & x = 5,6 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of k, $E(X)$ and $Var(X)$.

Use the fact that the total probability equals to 1 to find k.	$\begin{aligned} \int_1^3 2k\,dx + \sum_{x=5}^{6} k &= 1 \\ \int_1^3 2k\,dx + k + k &= 1 \\ [2kx]_1^3 + 2k = 6k - 2k + 2k \\ 6k = 1 \Rightarrow k &= \frac{1}{6} \end{aligned}$
Find the expectation for the first part where X is continuous using integration, then the expectation for the second part where X is discrete. Find the sum.	$\begin{aligned} \int_1^3 x\, f(x)\,dx &= \int_1^3 \frac{x}{3}dx = \left[\frac{x^2}{6}\right]_1^3 = \frac{9}{6} - \frac{1}{6} = \frac{4}{3} \\ \sum_{x=5}^{6} x\, f(x) &= 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{11}{6} \\ E(X) &= \frac{4}{3} + \frac{11}{6} \end{aligned}$
Find the variance of the continuous part of X.	$\begin{aligned} \int_1^3 x^2 f(x)\,dx &= \int_1^3 \frac{x^2}{3}dx = \left[\frac{x^3}{9}\right]_1^3 = \frac{27}{9} - \frac{1}{9} = \frac{26}{9} \\ \frac{26}{9} - \left(\frac{4}{3}\right)^2 &= \frac{26-16}{9} = \frac{10}{9} \end{aligned}$
Find the variance of the discrete part of X.	$\begin{aligned} \sum_{x=5}^{6} x^2 f(x) &= \frac{1}{6} \times 5^2 + \frac{1}{6} \times 6^2 = \frac{25+36}{6} = \frac{61}{6} \\ \frac{61}{6} - \left(\frac{11}{6}\right)^2 &= \frac{245}{36} \end{aligned}$
Find $Var(X)$ by adding up the variance of the two parts together.	$Var(X) = \frac{10}{9} + \frac{245}{36} = \frac{285}{36} = \frac{95}{12}$