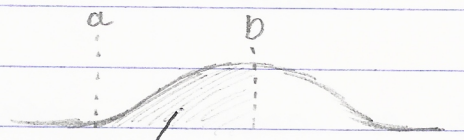
Continuous Random Variables

Fundamentals

- They can take on an infinite no. of possible values between an interval.
- Modelled by a curve, $f(x)$, called the probability <u>density</u> function (p.d.f).
- The probability is the area under the curve.



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

($P(X = a) = 0$ $\because$ it is an infinitely small area)

- Rules:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad , \quad f(x) \geq 0 \;\; \forall x \quad \because \text{ probability}$$

Cumulative distribution function (c.d.f)

- Has the symbol $F(x)$ where $F(a) = P(X \leq a)$ $= \int_{-\infty}^{a} f(x)\,dx \Rightarrow$

p.d.f $\xrightarrow{\text{integrate}}$ c.d.f

$f(x)$ $\xleftarrow{\text{differentiate}}$ $F(x)$

- Useful for medians, $F(x) = 0.5$, & quartiles, $F(x) = 0.25$.

Example

$$f(x) = \begin{cases} 1/4, & 0 \le x < 1 \\ x^3/5, & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow F(x) = \begin{cases} \frac{1}{4}x + C_1 & 0 \le x < 1 \\ x^4/20 + C_2 & 1 \le x \le 2 \\ 1 & \text{otherwise} \end{cases}$$

REMEMBER cumulative!

when $F(0) = 0$, you have no area (no probability so $f(x)$)

$\Rightarrow \quad 1/4\,(0) + C_1 \quad \Rightarrow \quad C_1 = 0$

when $x = 1$, $\quad 1/4\,x = x^4/20 + C_2 \Rightarrow 1/4\,x + 1/20 + C_2$

$\Rightarrow C_2 = 1/5$

$\because$ remember, it is cumulative

OR

$$\frac{x^4}{20} + C_2 = 1 \text{ when } x = 2 \Rightarrow C_1 = 4/20 = 1/5$$

Expectation, variance, & function

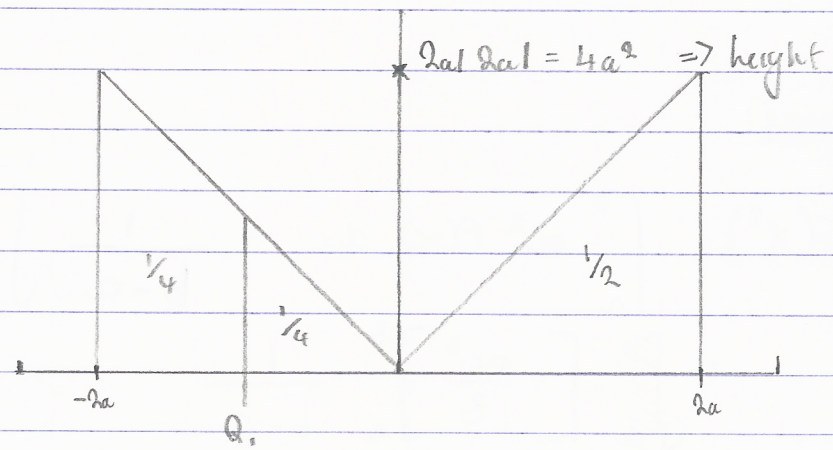
- $E(X) = \mu = \int x f(x)\,dx$ (mean) — each value · its probability
- $Var(X) = \sigma^2 = E(X^2) - \mu^2$ — $\int x^2 f(x)\,dx - \mu^2$
- For a function $E(g(X)) \Rightarrow \int g(x) f(x)\,dx$

Example 1

Sketch the pdf:

$$f(x) = \begin{cases} 2a\,|x| & -2a < x < 2a \\ 0 & \text{otherwise} \end{cases}$$

and find the CDF



Total area = $2a \times 4a^2 = 8a^3 = 1 \Rightarrow \underline{a = 1/2}$

$\because y = -x$

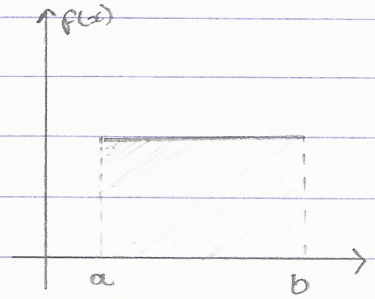
$$\int_{Q_1}^{0} (-x)\, dx = 1/4$$

$$\left[-\frac{x^2}{2} \right]_{Q_1}^{0} = 1/4 \Rightarrow \frac{Q_1^2}{2} = 1/4$$

$$\Rightarrow Q_1 = \pm\sqrt{1/2}$$

$$\Rightarrow Q_1 = -\sqrt{2}/2 \quad \because -2a < Q_1 < 0$$

Continuous Uniform (Rectangular) Distribution



Area = $(b-a)h = 1$

$\Rightarrow h = \frac{1}{b-a}$

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

<u>Mean:</u>

$$E(X) = \int_a^b x\, f(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \int_a^b x\, dx$$

$$= \frac{1}{b-a} \left[\frac{1}{2} x^2 \right]_a^b = \frac{1}{b-a} \left[\frac{b^2 - a^2}{2} \right]$$

$$= \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}} = \frac{b+a}{2}$$

Variance:

$$E(X^2) = \int_a^b x^2 f(x)\,dx = \frac{1}{b-a}\int_a^b x^2\,dx$$

$$= \left[\frac{x^3}{3}\right]_a^b \frac{1}{b-a}$$

$$= \frac{b^3-a^3}{3} \times \frac{1}{b-a}$$

$$= \frac{\cancel{(b-a)}(ab+a^2+b^2)}{3} \times \frac{1}{\cancel{b-a}}$$

$$= \frac{b^2+ab+a^2}{3}$$

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2$$

$$= \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{4b^2+4ab+4a^2}{12} - \frac{3a^2+6ab+3b^2}{12}$$

$$= \frac{a^2-2ab+b^2}{12}$$

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

CRVs, Integration technique using made-up example

$$f(x) = \begin{cases} 3x, & 2 \le x \le 5 \\ \frac{7x^3}{5}, & 5 \le x \le 10 \\ -2x, & 12 \le x \le 14 \\ 0, & \text{Otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{3x^2}{2} + c, & 2 \le x \le 5 \\ \frac{7x^4}{20} + d, & 5 \le x \le 10 \\ e, & 10 \le x \le 12 \\ -x^2 + f, & 12 \le x \le 14 \\ 1, & x \ge 14 \end{cases}$$

When $x = 2 \Rightarrow \frac{3(2)^{\cancel{2}}}{\cancel{2}} + c = 0$

$\Rightarrow c = -6$

$x = 5 \Rightarrow \frac{3(5)^2}{2} - 6 = \frac{7(5)^4}{20} + d$

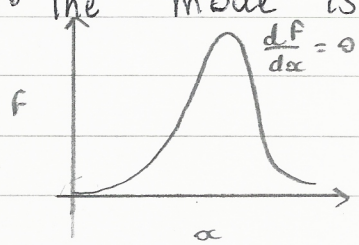
$\Rightarrow d = -187.25$

$x = 10 \Rightarrow \frac{7(10)^4}{20} - 187.25 = 3312.75 = e$

...

Extra Note

- The mode is where a pdf is highest



(whilst you cannot get this exact value, a lot of values will cluster here).