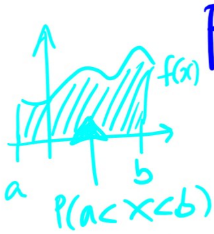
3Y03 - 3J04 PROBABILITY & STATISTICS FOR (CIVIL) ENGINEERING

C01 - Lecture 12

Last time CONTINUOUS RANDOM VARIABLES X

AND PROBABILITY DENSITY FUNCTIONS (pdf)

$f(x) \geq 0$ with $\int_{-\infty}^{\infty} f(x)\,dx = 1$ so that



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

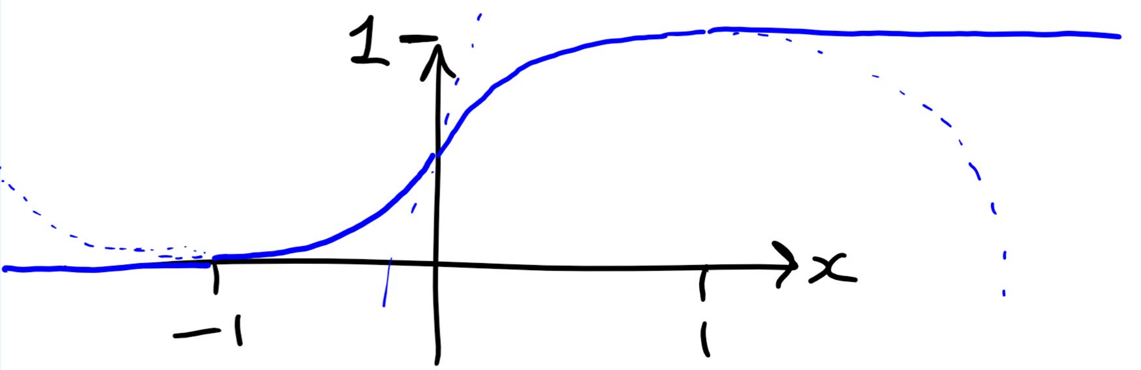
(Probability represented by area under graph of $f(x)$.)

$$P(X \leq x) = P(X < x) = \int_{-\infty}^{x} f(t)\,dt$$

This is $F(x)$, the cumulative distribution function; by definition this is the antiderivative of $f(x)$ i.e. $\frac{d}{dx} F(x) = f(x)$.

Example Suppose the cumulative distribution function of a random variable X is given by

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{1}{2}(x+1)^2, & -1 < x \leq 0 \\ 1 - \frac{(1-x)^2}{2}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$



continuous function not a step function (unlike the discrete case)

Find $P(X \le \frac{1}{2})$, $P(X > 0)$ and $f(x)$.

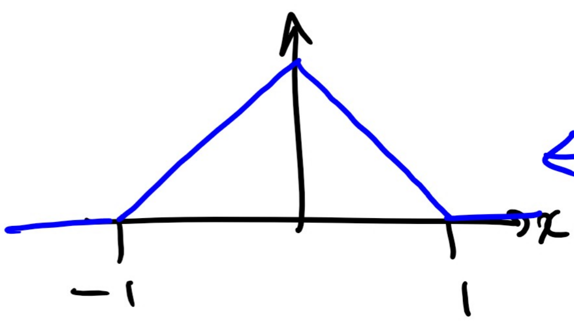
<u>Solution</u> $P(X \le 1/2) = F(\frac{1}{2}) = 1 - \dfrac{(1-\frac{1}{2})^2}{2}$

$$= 1 - \frac{1/4}{2} = \frac{7}{8}.$$

$$P(X > 0) = 1 - P(X \le 0) = 1 - F(0) = 1 - \tfrac{1}{2} = \tfrac{1}{2}.$$

$\frac{1}{2}(x+1)^2$ $\quad$ $1 - \frac{1}{2}(1-x)^2$

$$f(x) = \frac{d}{dx} F(x) = \begin{cases} 0 & \text{for } x \le -1 \\ x+1 & \text{for } -1 < x \le 0 \\ 1-x & \text{for } 0 < x \le 1 \\ 0 & \text{for } x \geq 1 \end{cases}$$



← A Triangular Distribution

<u>Example</u> For our earlier waiting time example we had

$$f(x) = \begin{cases} 0.5\, e^{-0.5x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find $F(x)$ and then $P(X > 3)$.

Solution

$$F(x) = \begin{cases} 0 & x < 0 \\ \displaystyle\int_0^x 0.5e^{-0.5t}\,dt & x \geq 0 \end{cases}$$

$$= \begin{cases} 0 & x < 0 \\ \left[-e^{-0.5t}\right]_0^x & x \geq 0 \end{cases}$$

$$= \begin{cases} 0 & x < 0 \\ 1 - e^{-0.5x} & x \geq 0. \end{cases}$$

$$P(X > 3) = 1 - P(X \leq 3) = 1 - F(3)$$

$$= 1 - \left(1 - e^{-0.5 \times 3}\right)$$

$$= e^{-1.5} = 0.2231.$$

Example The median is the "measure of central tendency" – the value m such that

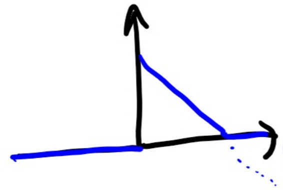
$$P(X \leq m) = P(X \geq m) = 0.5.$$

(If this defines a unique m; more general definition later in course.)

Let's take a continuous random variable X with pdf

$$f(x) = \begin{cases} 2(1-x) & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the median.



Solution Solve $P(X \le m) = 0.5$

i.e. $F(m) = 0.5$

$$0.5 = F(m) = \int_0^m 2(1-t)\,dt = \left[2t - t^2\right]_0^m$$

$$= 2m - m^2$$

$$\Rightarrow m^2 - 2m + 0.5 = 0 \Rightarrow m = 1 \pm \frac{1}{\sqrt{2}}$$

The only valid solution is in $[0,1]$ so

$$m = 1 - \frac{1}{\sqrt{2}} \approx 0.2929.$$

4.4 Mean & Variance of Continuous Random Variables

$\int$ replaces $\sum$

If X is a continuous random variable with pdf $f(x)$, then the mean is

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

(Centre of mass of pdf $f(x)$.)

And for any function of X, say $h(X)$, the <u>expected value of $h(X)$</u> $= E(h(X)) = \int_{-\infty}^{\infty} h(x) f(x)\,dx$

Special case: Variance $V(X) = \sigma^2$

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

$$= E(X^2) - (E(X))^2$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2.$$

<u>Example</u> The size X of particles of contamination is modelled with pdf

$$f(x) = \begin{cases} 2/x^3, & x > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find mean and variance.

<u>Solution</u> $\mu = E(X) = \int_1^{\infty} x f(x)\,dx = \int_1^{\infty} \frac{2}{x^2}\,dx$

$$= \left[-2x^{-1}\right]_1^{\infty} = 2.$$

$$V(X) = E(X^2) - \mu^2 = \int_1^\infty x^2 f(x)\,dx - 2^2$$

$$= \int_1^\infty \frac{2}{x}\,dx - 4 = \left[2\ln x\right]_1^\infty - 4$$

$$= \infty.$$

4.6 Normal Distribution

also known as Gaussian Distribution

- experiment repeated many times
- average / total results over all repetitions usually has normal distribution
- arises in many different areas e.g. velocity of molecules

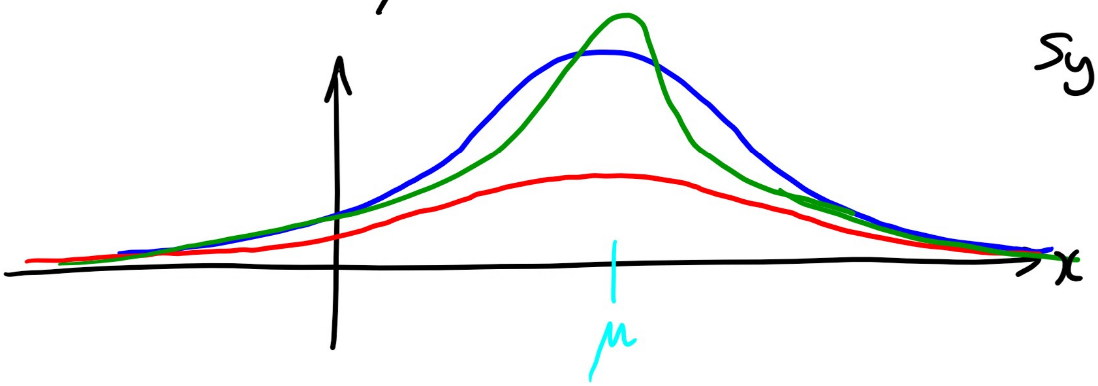
Rather strange pdf $f(x) = \dfrac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for all $x \in \mathbb{R}$

There are parameters μ, σ where $\mu \in \mathbb{R}$
$\sigma > 0$

There is a reason for the choice of letters:

if X has a normal distribution with parameters μ and σ (written $X \sim N(\mu, \sigma^2)$) then

$E(X) = \mu$ and $V(X) = \sigma^2$.



Symmetric &
the higher the
value of σ^2
the flatter
the curve