RANDOM VARIABLES

<u>Defn</u>: A random variable is a function that associates a real number with each element in the sample space.

<u>Notation</u>: Capital letter, say X, denotes a random variable and its corresponding small letter, x in this case, denotes one of its values.

<u>Example 1</u>: X: Number of defective products when 3 products are tested.

D: defective
N: non-defective

Outcomes in sample space	x : value of X
D D D	3
D D N	2
D N D	2
D N N	1
N D D	2
N D N	1
N N D	1
N N N	0

If the sample space contains a finite number of possibilities such as this case, it is called a <u>discrete sample space</u>.

<u>Example 2</u>: Y: Number of products tested before a defective product is found.

Sample Space	y
D	1
N D	2
N N D	3
N N N D	4
N N N N D	5
N N N N N D	6
⋮	⋮

If the sample space contains an unending sequence with as many elements as there are integers (whole numbers), it is still called a discrete ~~ex~~ sample space.

This is an example of such a case.

Example 1 and 2 are discrete random variables.

<u>Example 3</u> : Z: Height in inches of a random person.

We can measure z to any degree of accuracy:

71 inches
71.2 inches
71.24
71.243

} Continuous sample space
Continuous random variable

If a sample space contains an infinite number of outcomes equal to the number of points on a line segment, it is called a <u>continuous sample space</u>.

<u>Example 4</u> : X: Power consumption of my house measured in kilowatts (kw) at a random time.

This is another example of a continuous random variable.

<u>Example</u> : Let X be the random variable : number of heads in 3 tosses of a fair coin

Sample space	x	$P(X=x)$
TTT	0	$P(X=0) = 1/8$
TTH	1	$P(X=1) = 3/8$
THT	1	$P(X=2) = 3/8$
THH	2	$P(X=3) = 1/8$
HTT	1	
HTH	2	
HHT	2	
HHH	3	

$P(X=x)$ → Notation: Probability that outcome is a specific x value

Probability function

Terminology for discrete random variables ⟹ Probability mass function
Probability distribution

<u>Defn</u>: $f(x)$ is a probability mass function for the discrete random variable X if, for each possible outcome

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$
3. $P(X = x) = f(x)$

<u>Example</u>: If the random variable X has n possible outcomes $x_1, x_2, \ldots, x_n$ then the above rules become

1. $f(x_i) \geq 0$ for $i = 1$ to n
2. $\sum_{i=1}^{n} f(x_i) = 1$
3. $P(X = x_i) = f(x_i)$

<u>Example</u>: X : Number of bits equal to 1 in an 8-bit random binary number.

The sample space has $2^8 = 256$ possible binary numbers. Each are equally likely. The random variable can have values between 0 and 8. Lets compute the probabilities

$P(X=0) = 1/256$ only 1 way to have all bits OFF

$P(X=1) = 8/256$ $\quad {}_8C_1 = \frac{8!}{7!\,1!} = 8$ ways

$\vdots$

$P(X=k) = \frac{{}_8C_k}{256}$

$\vdots$

$P(X=8) = 1/256$

$$f(x) = \frac{{}_8C_x}{256}$$

<u>Notice</u>: 1. $f(x) \geq 0$ for $0 \leq x \leq 8$

2. $\sum_x f(x) = \sum_{k=0}^{8} \frac{{}_8C_k}{256} = \frac{2^8}{256} = 1$

3. $P(X = x) = f(x)$

Example continued.

What is the probability that there are less than or equal to 3 bits ON?

$$\underbrace{P(X \le 3)}_{\text{this is like an event}} = \underbrace{\sum_{k=0}^{3} P(X=k)}_{\text{These are the probabilities of outcomes in the event.}} = \sum_{k=0}^{3} \frac{{}_8C_k}{256}$$

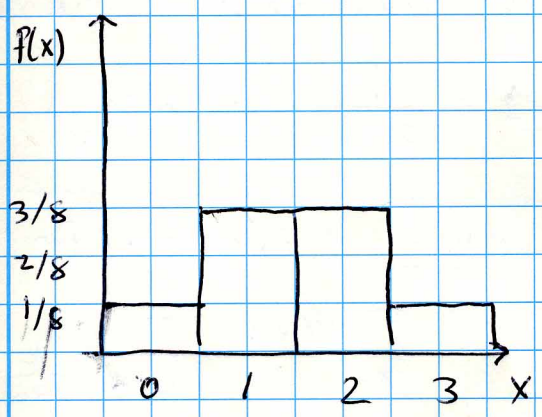
<u>Defn</u>: The cumulative distribution function $F(x)$ of a discrete random variable X with a probability mass function $f(x)$ is

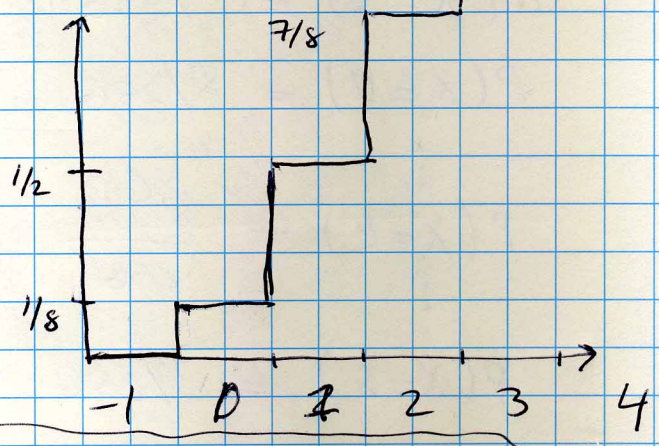
$$F(x) = P(X \le x) = \sum_{t \le x} f(t)$$

<u>Example</u>: Number of heads in 3 tosses of a fair coin.

$f(0) = 1/8 \quad f(1) = 3/8 \quad f(2) = 3/8 \quad f(3) = 1/8$

$F(0) = 1/8 \quad F(1) = 4/8 \quad F(2) = 7/8 \quad F(3) = 8/8$





* $F(x)$ is monotonically increasing since $f(x) \ge 0$.
* $F(-\infty) = 0$ * $F(\infty) = 1$

Continuous Probability Distributions

A continuous random variable has a probability of zero of assuming exactly any one of its values!

<u>Example</u>: height of random person

$$P(X = 164\text{ cm}) = 0 \quad !!!$$

<u>Why?</u> Between any two values, say 163.5 cm and 164.5 cm, or even 163.999 and 164.001, there are any infinite number of possible height (x) values. The probability of $x = 164.00000\ldots$ (to arbitrary precision) is 0.

<u>What to do?</u> With continuous random variables we talk about the probability of x being in some interval, like $P(a<X<b)$, rather than x assuming a precise value like $P(X=a)$.

<u>Note</u>: $P(a < X \leq b) = P(a < X < b) + \cancel{P(X=b)}$ 0

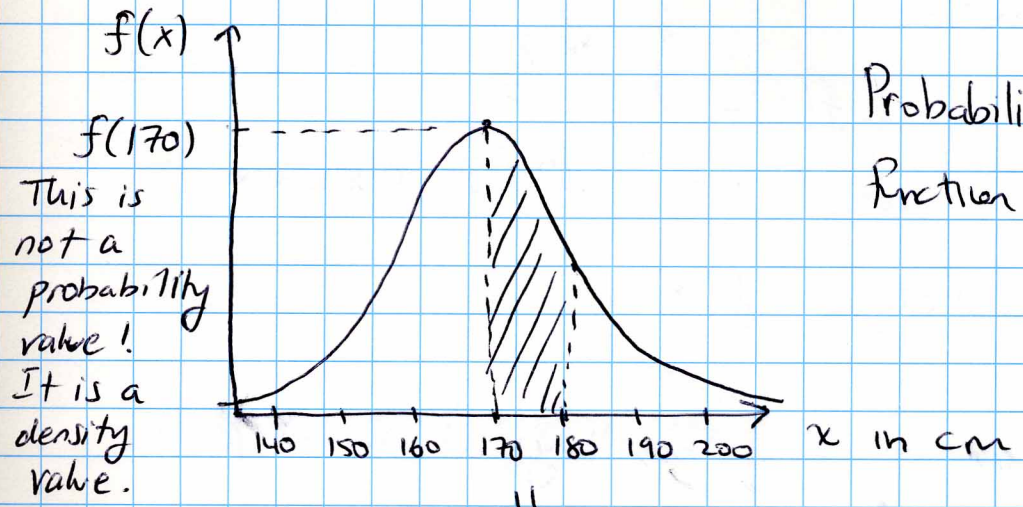
This holds <u>only</u> for continuous random variables.

<u>Defn.</u> The function $f(x)$ is a probability density function for the continuous random variable X if

1. $f(x) \geq 0$ for all $x \in \mathbb{R}$ (all real numbers)
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $P(a < X < b) = \int_a^b f(x)\,dx$

<u>Defn</u>: The cumulative distribution function $F(x)$ of a continuous random variable X with density function $f(x)$ is $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

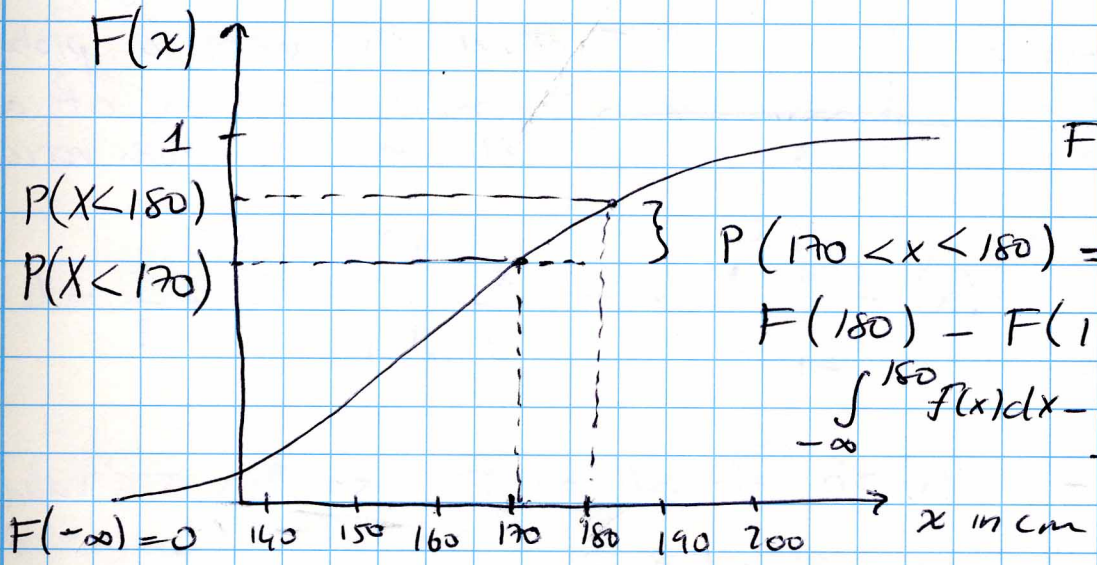
<u>Example</u>: Height of a random person over 21



Probability density function

$\Downarrow$

Shaded area is $P(170 < x < 180) = \int_{170}^{180} f(x)\,dx$



$F(\infty) = 1$

$P(170 < x < 180) = F(180) - F(170) = \int_{-\infty}^{180} f(x)\,dx - \int_{-\infty}^{170} f(x)\,dx$

* $P(a < X < b) = F(b) - F(a)$
* From the definition of $F(x)$ we have

$$f(x) = \frac{dF(x)}{dx}$$

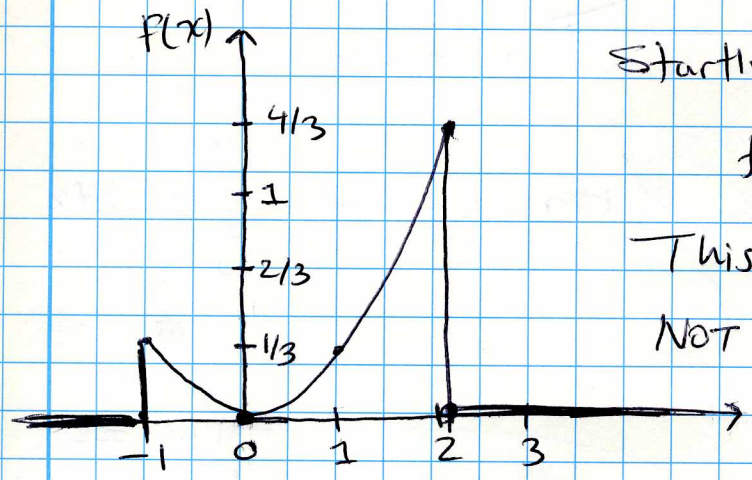
$f(x)$ is the derivative of $F(x)$

* $P(X = a) = \int_a^a f(x)\,dx = F(a) - F(a) = 0$

<u>Example</u> : A random variable X has the probability density function

$$f(x) = \begin{cases} x^2/3 , & -1 < x \leq 2 \\ 0 , & \text{elsewhere} \end{cases}$$

a) Lets sketch $f(x)$



Startling (at first) fact:

$f(2) = 4/3 > 1$.

This ok for a density function.

NOT ok. if this were a probability mass function of a discrete random variable.

b) $$\int_{-\infty}^{\infty} f(x) = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1$$

Also $f(x) \geq 0$ for all $x \in \mathbb{R}$ so this is a valid probability density function.

c) What is the probability that X is between 0 and 1?

Directly from density function:

$$P(0 < X < 1) = \int_0^1 \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_0^1 = \frac{1}{9}$$

Via the cumulative distribution function $F(x)$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

<u>Case 1</u> $x < -1$ $\quad F(x) = \int_{-\infty}^{x} 0\,dt = 0$

<u>Case 2</u> $-1 \le x \le 2$

$$F(x) = \underbrace{\cancel{\int_{-\infty}^{-1} 0\,dt}}_{0} + \int_{-1}^{x} \frac{t^2}{3}\,dt$$

$$= \frac{t^3}{9}\Big|_{-1}^{x} = \frac{x^3+1}{9}$$

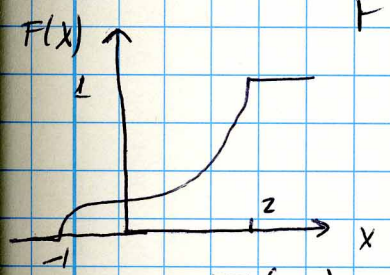
<u>Case 3</u> $x > 2$

$$F(x) = \cancel{\int_{-\infty}^{-1} 0\,dt} + \int_{-1}^{2} \frac{t^2}{3}\,dt + \cancel{\int_{2}^{x} 0\,dt}$$

$$= \frac{t^3}{9}\Big|_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1$$

<u>So</u>

$$F(x) = \begin{cases} 0, & x < -1 \\ \dfrac{x^3+1}{9}, & -1 \le x \le 2 \\ 1, & x > 2 \end{cases}$$



* $F(x)$ is monotonically increasing always.

Example 3.21 from textbook 3.5

Exercise 3.5 (textbook)

a) $f(x) = c(x^2+4) \quad x = 0,1,2,3$

Find the value of c that makes $f(x)$ a valid probability mass function (discrete)

Soln: We require $\sum_{x=0}^{3} c(x^2+4) = 1$

$$4c + 5c + 8c + 13c = 1$$

$$30c = 1 \implies c = 1/30$$

b) $f(x) = c\binom{2}{x}\binom{3}{3-x}$ for $x = 0,1,2$

Again $\sum_{x=0}^{2} f(x) = 1$

$$c\left[\binom{2}{0}\binom{3}{3} + \binom{2}{1}\binom{3}{2} + \binom{2}{2}\binom{3}{1}\right] = 1$$

$$c[1 + 6 + 3] = 1$$

$$10c = 1 \implies c = 1/10$$

Exercise 3.6 (textbook)

$$f(x) = \begin{cases} \dfrac{20000}{(x+100)^3}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Shelf life of medicine.
Continuous r.v.

a) $P(X > 200) = ?$

$$P(X > 200) = \int_{200}^{\infty} \frac{20000\,dx}{(x+100)^3} = -\frac{10000}{(x+100)^2}\Bigg|_{200}^{\infty}$$

$$= -\cancel{\frac{10000}{\infty}}^{\,0} + \frac{10000}{300^2} = 1/9$$

b) $P(80 < X < 200) = \displaystyle\int_{80}^{200} \frac{20000}{(x+100)^3}\,dx$

$$= \frac{-10000}{(x+100)^2}\Bigg|_{80}^{120} = 0.102$$

* $P(X \leq 200) = ?$

$$P(X \leq 200) = 1 - P(X > 200)$$

$$= 1 - 1/9 = 8/9$$

Exercise 3.12 (textbook)

$$F(t) = \begin{cases} 0, & t < 1 \\ 1/4, & 1 \le t < 3 \\ 1/2, & 3 \le t < 5 \\ 3/4, & 5 \le t < 7 \\ 1, & t \ge 7 \end{cases}$$

Notice a) this is a discrete random variable because t is an <u>integer</u> number of years and b) The cumulative dist. function is given.

a) $P(T=5) = F(5) - F(4) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

b) $P(T > 3) = 1 - P(T \le 3)$
$= 1 - F(3) = 1 - 1/2 = 1/2$

c) $P(1.4 < T < 6) = F(6) - F(1.4)$
$= \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

$F(1.4) = 1/4$ because $F(t)$ is like a step function. For all t values between 1 and 3, it takes on the value $1/4$.

Exercise 3.21 (textbook)

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Continuous prob. density fn.

a) Evaluate k.

$$\int_{-\infty}^{\infty} f(x)dx = 1$$ since $f(x)$ is a probability density function

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 k\sqrt{x}\,dx = \frac{2k}{3}x^{3/2}\Big|_0^1$$

$$= \frac{2k}{3} = 1$$

Therefore $k = 3/2$

b) $F(x) = \int_{-\infty}^{x} f(t)dt$

for $x < 0$ $F(x) = 0$

for $x > 1$ $F(x) = 1$

for $0 < x < 1$ $F(x) = \int_0^x \frac{3\sqrt{t}}{2}dt = t^{3/2}\Big|_0^x = x^{3/2}$

$$F(x) = \begin{cases} 0, & x < 0 \\ x^{3/2}, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$$

$P(0.3 < X < 0.6) = F(0.6) - F(0.3) = 0.3004$