Lecture 11

The Basic Numerical Quantities Associated to a Continuous X

In this lecture we will introduce four basic numerical quantities associated to a continuous random variable X. You will be asked to calculate these (and the cdf of X) given $f(x)$ on the midterms and the final.

These quantities are

1. The p-th percentile $\eta(p)$
2. The α-th critical value x_α.
3. The expected value $E(X)$ or μ.
4. The variance $V(X)$ or σ^2.

I will compute all these for the linear distribution and $U(a,b)$.

Percentiles and Critical Values of Continuous Random Variables

Percentiles

Greek letter eta

Let p be a number between 0 and 1. The 100p-th percentile, denoted $\eta(p)$, of a continuous random variable X is the unique number satisfying

$$P(X \le \eta(p)) = p \qquad (\#)$$

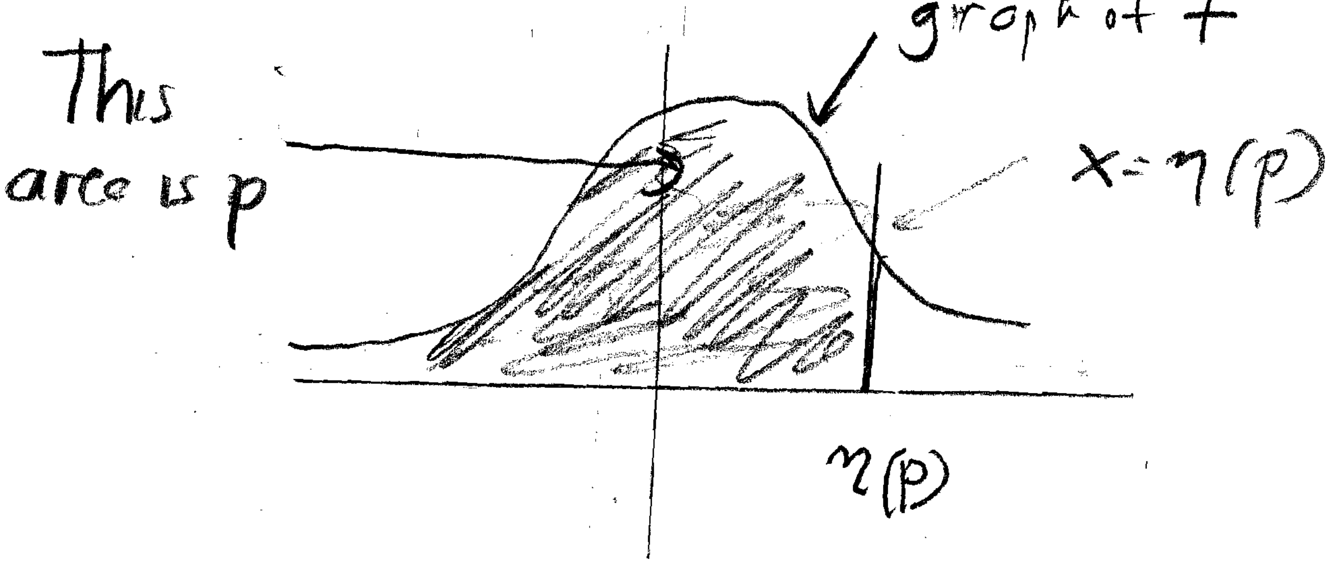
or

$$F(\eta(p)) = p \qquad (\#\#)$$

So if you know F you can find $\eta(p)$. Roughly

$$\eta(p) = F^{-1}(p)$$

The geometric interpretation of $\eta(p)$ is very important



<u>The geometric interpretation of (#)</u>

$\eta(p)$ is the number such that the vertical line $x = \eta(p)$ cuts off area p to the left under the graph of $f(x)$.

(this is the picture above)

Special Case   The median $\tilde{\mu}$

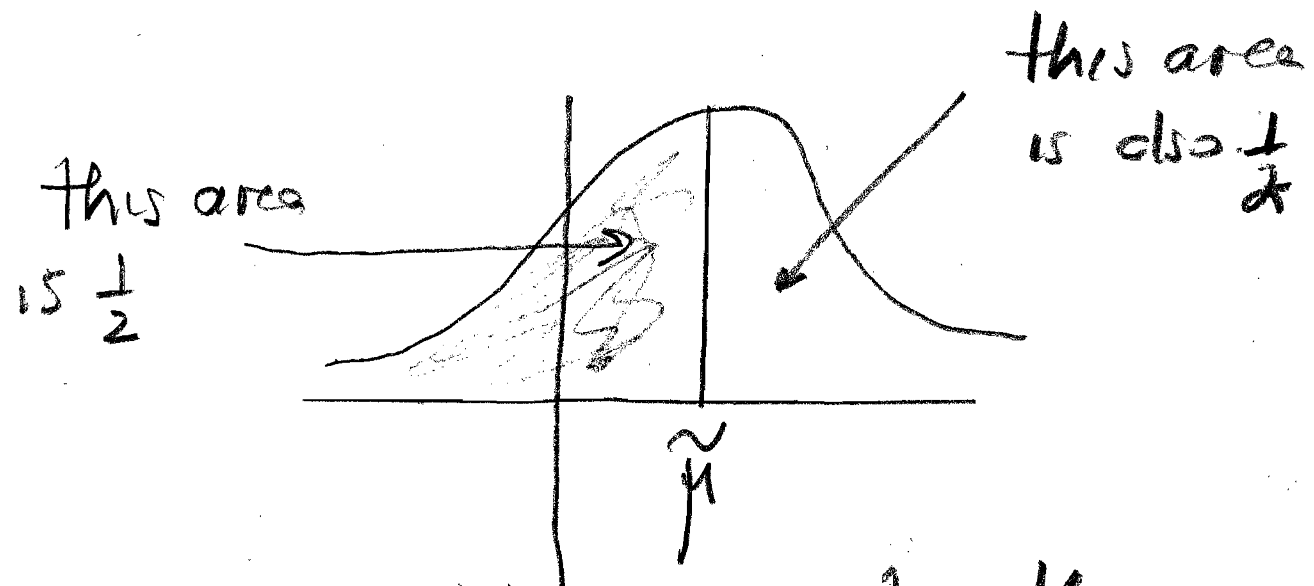
The median $\tilde{\mu}$ is the unique number so that

$$P(X \le \tilde{\mu}) = \frac{1}{2}$$

or $$F(\tilde{\mu}) = \frac{1}{2}$$

so the median is the 50-th percentile

The picture



Since the total area is 1, the area to the right of the vertical line $x = \tilde{\mu}$ is also $\frac{1}{2}$. So $x = \tilde{\mu}$ bisects the area.

Critical Values

Roughly speaking if you switch left to right in the definition of percentile you get the definition of the critical value. Critical values play a key role in the formulas for confidence intervals (later).

Definition Let α be a real number between 0 and 1. Then the α-th critical value, denoted x_α, is the unique number satisfying

$$P(X \geq x_\alpha) = \alpha \quad (b)$$

Let's rewrite (b) in terms of F. We have

$$P(X \geq x_\alpha) = 1 - P(X \leq x_\alpha)$$
$$= 1 - F(x_\alpha)$$

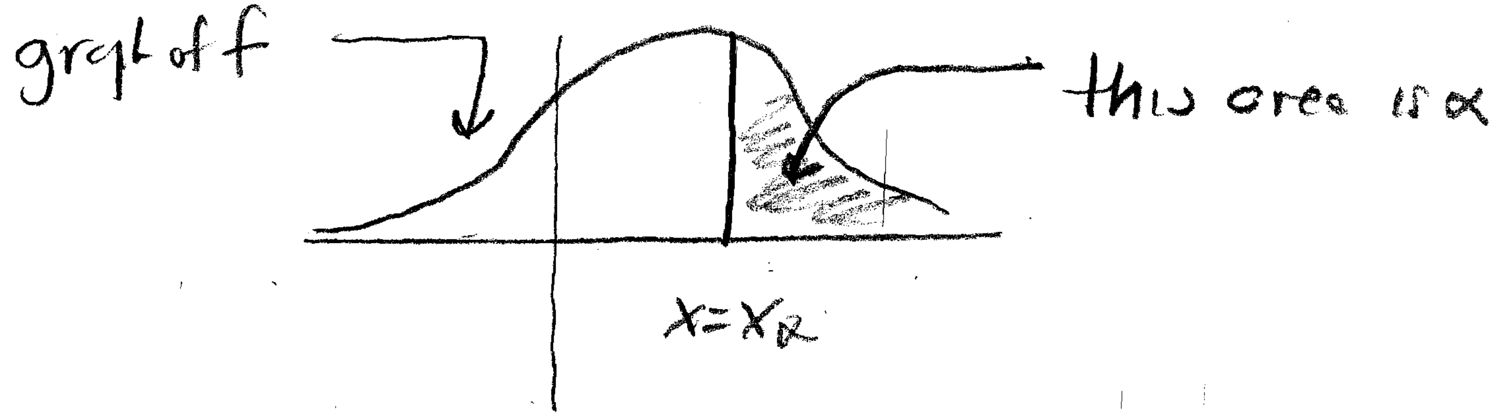
So (b) becomes

$$1 - F(x_\alpha) = \alpha$$

$$F(x_\alpha) = 1 - \alpha$$
$$x_\alpha = F^{-1}(1-\alpha) \quad \text{(bb)}$$

What about the geometric interpretation?

The geometric interpretation



x_α is the number so that the vertical line $x = x_\alpha$ cuts off area α to the right under the graph of $f(x)$.

Relation between critical values and percentiles

$x = x_\alpha$ cuts off area $1-\alpha$ to the left since the total area is 1.

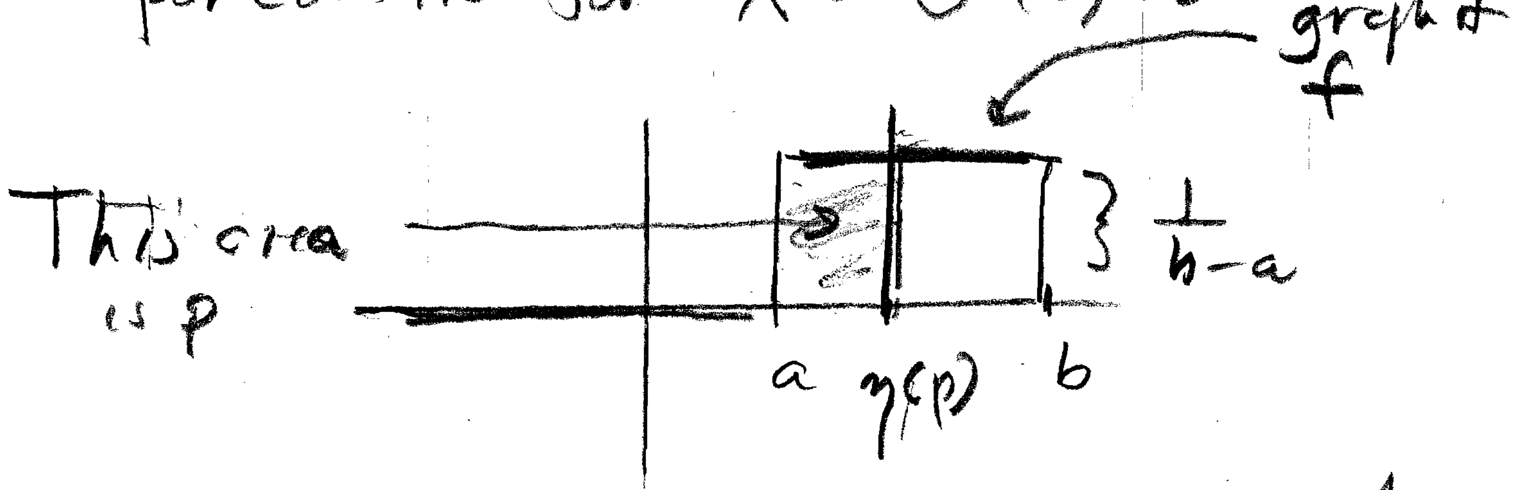
But $\eta(1-\alpha)$ is the number such that $x = \eta(1-\alpha)$ cuts off area $1-\alpha$ to the left.

So $$x_\alpha = \eta(1-\alpha)$$

Computation of Examples

Example 1 $X \sim U(a, b)$

Lets compute the $\eta(p)$-th percentile for $X \sim U(a, b)$



So the point $\eta(p)$ between a and b must have the property that the area of the shaded box is p. But the base of the box is $\eta(p) - a$ and the height is $\frac{1}{b-a}$ so

$$\text{Area} = bh = (\eta(p) - a)\left(\frac{1}{b-a}\right)$$ so

$$(\eta(p) - a)\left(\frac{1}{b-a}\right) = p$$ or

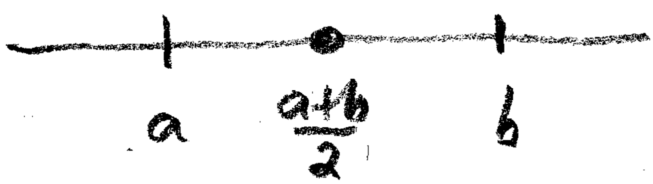
$$\eta(p) = a + p(b-a) = (1-p)a + pb \qquad (*)$$

How about the median $\tilde{\mu}$.
So we want $\eta(\frac{1}{2})$. By (*) we have

$$\tilde{\mu} = \eta(\tfrac{1}{2}) = a + \frac{b-a}{2} = \frac{a+b}{2}$$

Remark

$\frac{a+b}{2}$ is the midpoint of the interval $[a,b]$



Critical Values for $U(a,b)$

$$x_\alpha = \eta(1-\alpha) = a + (1-\alpha)(b-a)$$

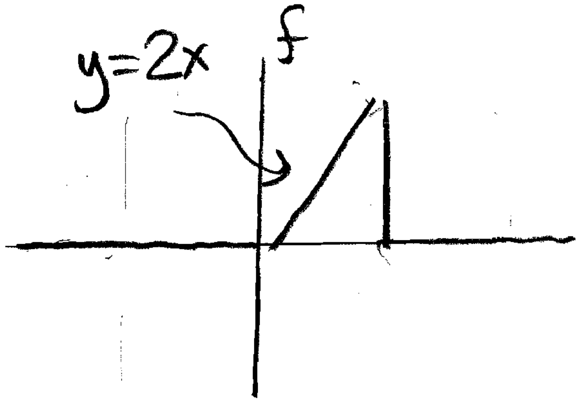
$$= a + b - a - \alpha b + \alpha a$$

So
$$x_\alpha = \alpha a + (1-\alpha) b.$$

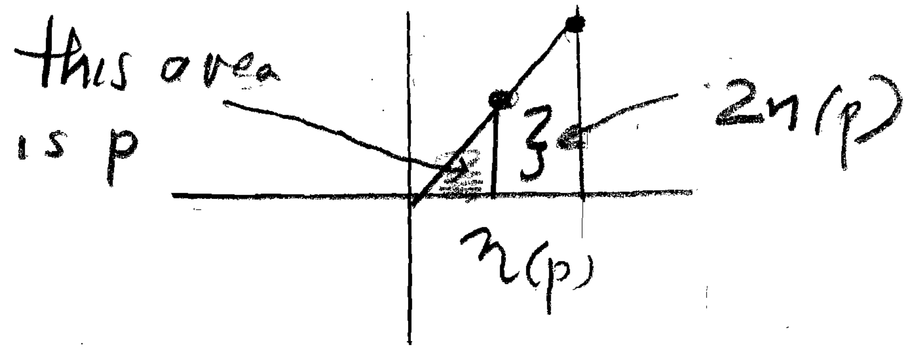
Example 2 The linear distribution

Recall the linear distribution has density

$$f(x) = \begin{cases} 0, & x < 0 \\ 2x, & 0 \le x \le 1 \\ 0, & x > 1 \end{cases}$$



The 100p-th percentile



We want the area of the triangle to be p. But the base is $\eta(p)$ and the height is $2\eta(p)$ so

$$A = \frac{1}{2} b h = \frac{1}{2} \eta(p)\,(2\eta(p))$$
$$= \eta(p)^2$$

We have to solve

$$\eta(p)^2 = p$$

so $\eta(p) = \sqrt{p}$

In particular

$$\tilde{\mu} = \eta\left(\tfrac{1}{2}\right) = \sqrt{\tfrac{1}{2}} = \frac{\sqrt{2}}{2}$$

This will be important later

Expected Value

Definition

The expected value or mean $E(X)$ or μ of a continuous random variable is defined by

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

We will compute some examples

Example 1

$X \sim U(a, b)$

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

$$= \frac{1}{b-a} \left(\frac{x^2}{2}\right)\Bigg|_{x=a}^{x=b} = \frac{1}{2} \frac{(b^2-a^2)}{b-a} = \frac{b+a}{2}$$

Now we showed on page 9 that if $X \sim U(a, b)$ then the median $\tilde{\mu}$ was given by

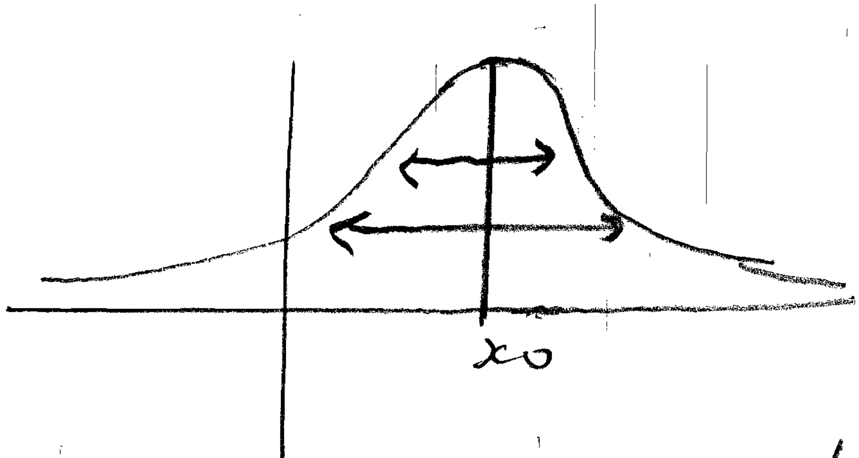
$$\tilde{\mu} = \frac{a+b}{2}$$

Hence in this <u>the mean is equal to the median</u>

$$\mu = \tilde{\mu} = \frac{a+b}{2}$$

$\geq$ This is not always the case as we will see shortly.

The "reason" $\mu = \tilde{\mu}$ is that $f(x)$ has a point of symmetry ie a point x_0 so that $f(x_0+y) = f(x_0-y)$



This means ~~that the graph~~ is symmetrical about the vertical line (mirror) $x = x_0$

<u>Proposition (Useful fact)</u>

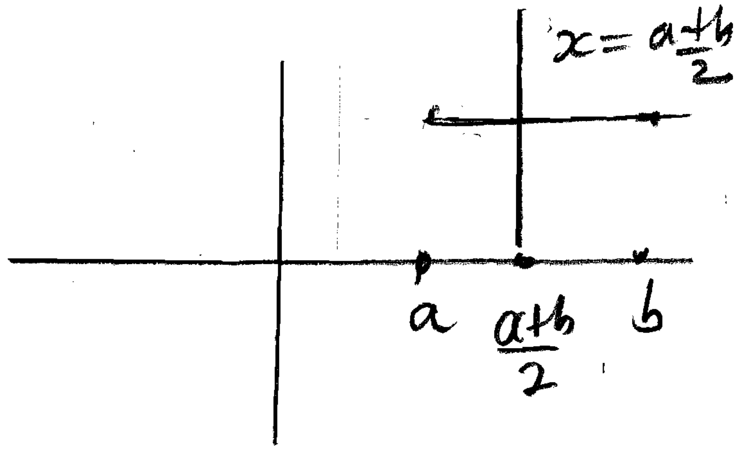
If x_0 is a point of symmetry for $f(x)$ then

$$\mu = \tilde{\mu} = x_0$$

Now if $X \sim U(a, b)$ then

$x_0 = \frac{a+b}{2}$ is a point of symmetry

for $f(x)$



For a change we will prove

the proposition.

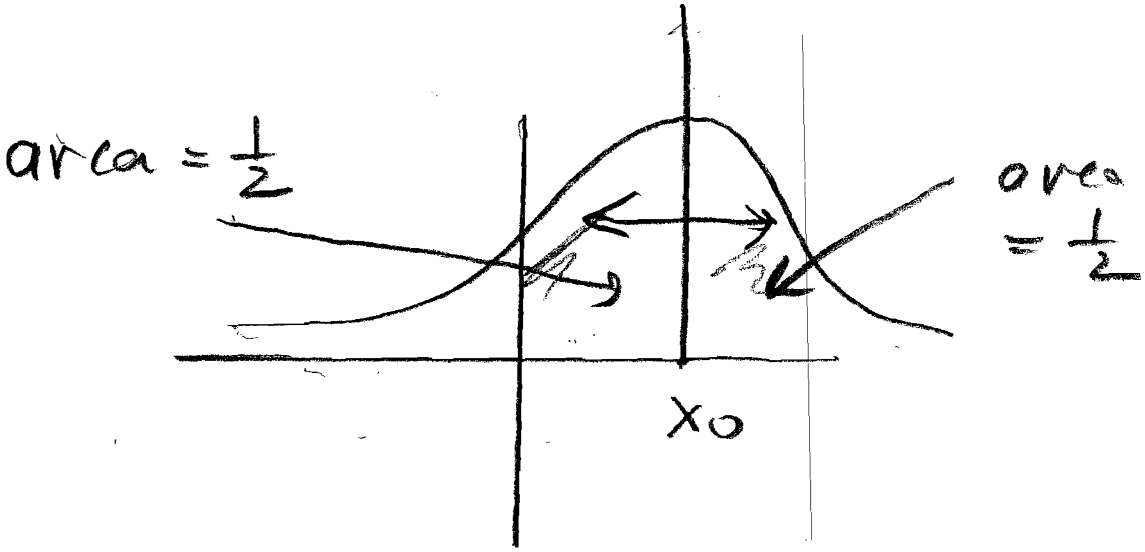
<u>Proof</u>

$\tilde{\mu} = x_0$ is immediate because

by symmetry there is equal

area to the left and right

of x_0.



Since the total area is 1 the area to the left of x_0 is $\frac{1}{2}$.
Hence $\tilde{\mu} = x_0$.

It is harder to prove

$$E(X) = \int_{-\infty}^{\infty} x f(x) = x_0$$

Trick: Since x_0 is a constant and $\int_{-\infty}^{\infty} f(x)\,dx = 1$ we have

$$\int_{-\infty}^{\infty} x_0 f(x)\,dx = x_0$$

Thus to show

$$\int_{-\infty}^{\infty} x f(x)\,dx = x_0$$

it suffices to show

$$\int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{\infty} x_0 f(x)\,dx$$

or

$$\int_{-\infty}^{\infty} (x-x_0) f(x)\,dx = 0$$

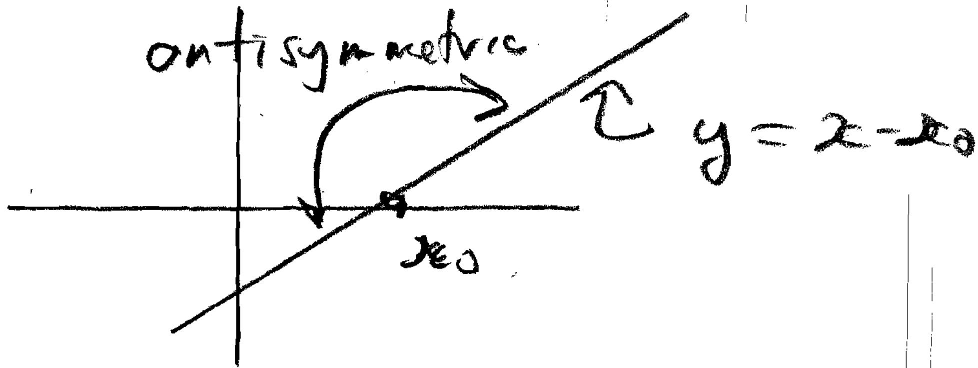
But if we put

$$g(x) = (x-x_0) f(x) \text{ then}$$

$g(x)$ is antisymmetric or "odd" about x_0

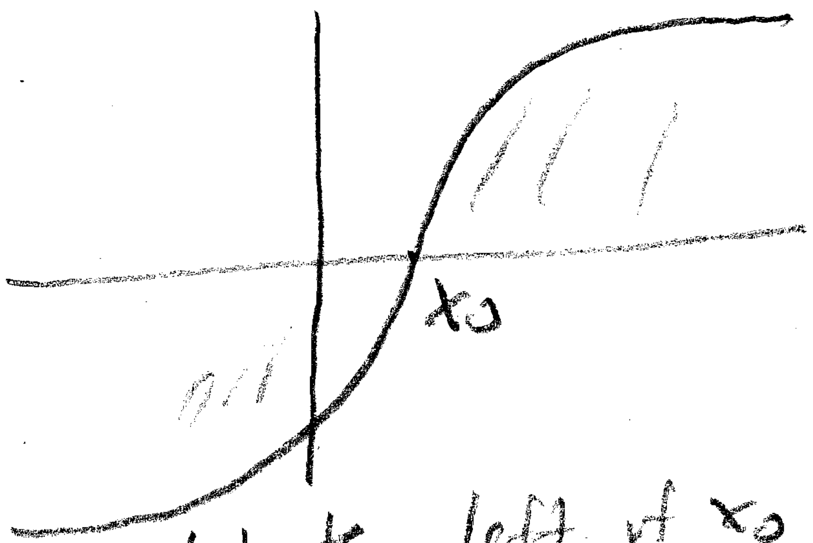
$$g(x_0+y) = -\,g(x_0+y)$$

This is because $x - x_0$ is



But antisymmetric · symmetric
= antisymmetric
(or odd · even = odd)

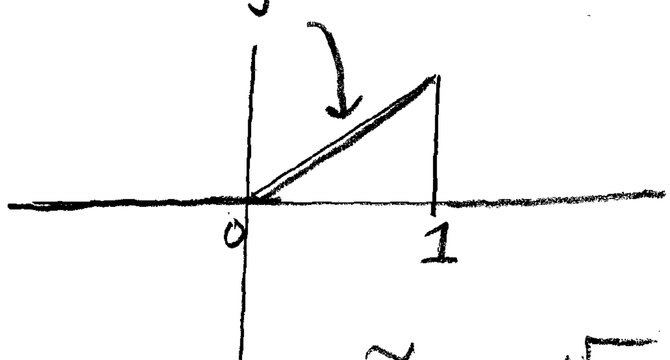
Finally the integral of an antisymmetric (or "odd") function from $-\infty$ to ∞ is zero



The integral to the left of x_0 cancels the area to the right $\square$

This fact can save a lot of painful computation of expected values.

Example 2 The linear distribution



We have seen $\tilde{\mu} = \frac{\sqrt{2}}{2}$, page 12.

$f(x)$ is certainly not symmetric so it is possible $\mu = \tilde{\mu}$ and we will see that it is the case

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

$$= \int_0^1 x(2x)\, dx$$

$$= 2\int_0^1 x^2\, dx$$

$$= 2\left(\tfrac{1}{3}\right) = \frac{2}{3}$$

Handy fact $\int_0^1 x^n = \frac{1}{n}$

So $\mu = 2/3$ and $\tilde{\mu} = \frac{2}{\sqrt{2}}$

They aren't equal, which one is bigger?

Variance

The Variance $V(X)$ or σ^2 of a continuous random variable is defined by

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

Remark Once we learn about change of continuous random variable we will see this is $E((X-\mu)^2)$

new random variable obtained from X using $h(x) = (x-\mu)^2$.

Once again there is a shortcut formula for $V(X)$.

<u>Proposition</u> <u>Shortcut Formula</u>

$$V(X) = E(X^2) - (E(X))^2$$
$$= E(X^2) - \mu^2$$

<u>This is the formula to use.</u>

<u>Example 1</u> <u>$X \sim U(a,b)$</u>

We know $\mu = \dfrac{a+b}{2}$. We have to compute $E(X^2)$

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

$$= \int_a^b x^2 \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a}\left(\frac{x^3}{3}\right)\Bigg|_{x=a}^{x=b}$$

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

So

$$\mu^2$$

$$V(X) = \frac{1}{3}(a^2+ab+b^2) - \left(\frac{a+b}{2}\right)^2$$

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

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

Example 2 The linear distribution

We have seen (pg 21)

$$\mu = \frac{2}{3}$$

We need $E(X^2)$.

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

$$= \int_0^1 x^2 (2x)\, dx$$

$$= 2 \int_0^1 x^3\, dx = 2\left(\frac{1}{4}\right) = \frac{1}{2}$$

So

$$V(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9}$$

$$= \frac{9}{18} - \frac{8}{18} = \frac{1}{18}$$