Primer on integration

Integration can be seen in two ways, though in the end they both come to the same thing (most of the time);

1) Finding the area under a curve
2) The reverse of differentiation.

We will start by dealing with 2). Although for our purposes we are more interested in 1), 2) shows us how we can actually find the integrals of a lot of common functions.

Let f(x) be a function. Then the indefinite integral of f(x), written $\int f(x)dx$ is a function F(x) such that $\frac{dF}{dx}$=f(x).

Note that, since the differential of any constant is zero, we can always add any constant to our indefinite integral.

For example, let f(x)=2x. We know that if F(x)=x^2, then $\frac{dF}{dx}$=2x. But this will also be true if F(x)=x^2+K for any constant K. So we write

$\int 2xdX$ =x^2+K

The constant K is called the constant of integration

Some rules for integration

You may verify these by considering the reverse process of differentiation.

1) If f(x)=a, a constant, then $\int f(x)dx$ = ax+K
2) If f(x) = x^n, with n≠-1, then $\int f(x)dx$ = x^{n+1}/(n-1)+K
3) If f(x) = 1/x, then $\int f(x)dx$=Ln(x)+K
4) If f(x) = e^x then $\int f(x)dx$ = e^x+K
5) For any constant a, $\int af(x)dx = a\int f(x)dx$
6) $\int f(x) + g(x)dx = \int f(x)dx + \int g(x)dx$

7) (Integration by parts) $\int f(x)\frac{dg}{dx}dx = f(x)g(x) - \int g(x)\frac{df}{dx}\,dx$

This last one is quite complicated: if we have a product, we must express one of the terms as the differential of some other function, and hope that this will simplify things.

<u>Example</u>

Find $\int xe^x dx$

Let f(x)=x, let dg/dx=e^x

Then df/dx=1, and g(x)=e^x (see rule 4)

So the answer is f(x)g(x)-$\int g(x)\frac{df}{dx}\,dx$

= xe^x - $\int 1.e^x dx$

= $xe^x - e^x$ + K.

Check: Let f(x) = $xe^x - e^x$

Then $\frac{df}{dx} = xe^x + e^x.1 - e^x$

(Using the product rule for differentiation)

= xe^x, as required.

8) $\int \frac{df}{dx}g(f(x))dx = G(f(x))$ where G(x)=$\int g(x)dx$ (substitution rule)

Example: find $\int 2x.e^{x^2}$ dx

Let f(x) = x^2, and g(x)=e^x. So df/dx=2x, and the function to be integrated is in the required form.

Then $G(x)=e^x$, so the answer to the integral is $G(f(x))=G(x^2) = e^{x^2}$. (Plus a constant of integration).

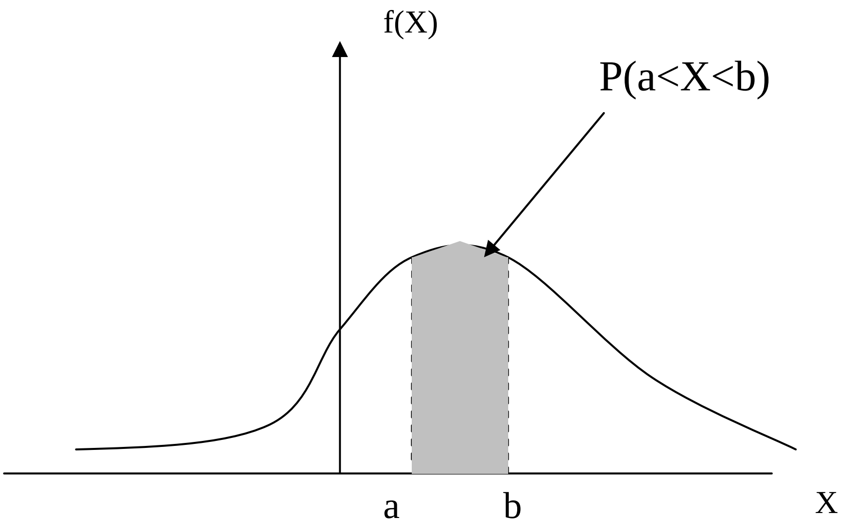
Check: Let $F(x)= e^{x^2}$. This can be written as $G(H(x))$ where $G(y)=e^y$, and $H(x)=x^2$. Then $dG/dy = e^y$, and $dH/dx=2x$

Then using the chain rule for differentiation, $\frac{dF}{dX} = \frac{dG}{dX}(H(x)).\frac{dH}{dx}$

$= e^{x^2}.2x$, as required.

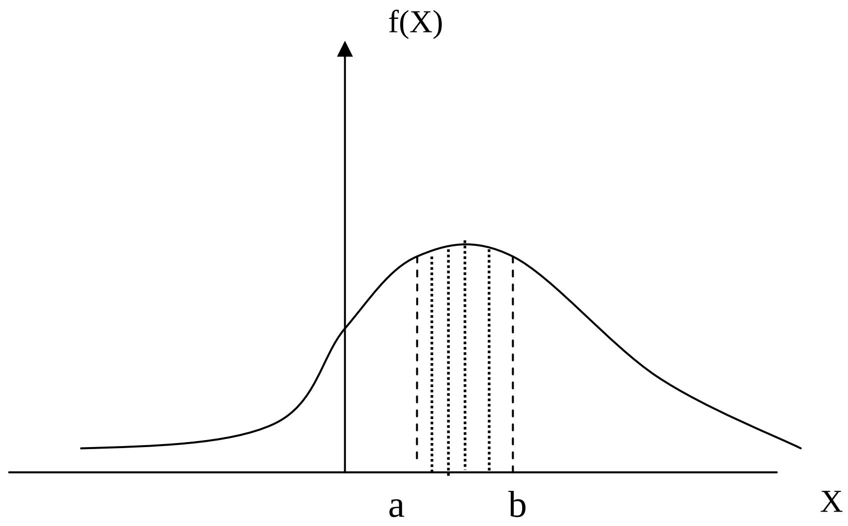
Definite integrals – finding the area under a curve

In statistics, we frequently need to find areas under a curve, and perform related integration. Namely, when we have a probability density function f(X) for a random variable X, the probability that X lies between two values a and b is defined to be the area under the curve of f(X) lying between X=a and X=b. (See below)



How do we find the area under a curve? If the curve consists of straight lines, it is relatively easy to solve by geometric means. With a curve, the idea is to split the area up into narrow vertical strips, and find the area of a rectangle in each strip. This gives an approximation to the total area. The narrower the strips, the closer the approximation. We define the

<u>definite integral of f(x) between a and b</u>, $\int_{X=a}^{X=b} f(X)dX$ to be the <u>limit</u> of these approximations as the strips get narrower. (See diagram below)



What is the meaning of the 'dX' in the integral? It means, essentially, the *width of a small strip of area under the curve*, that is, a small change in X over which we are measuring the area of a strip. ('d' in maths usually refers to a change.)

When we calculate our approximate areas, the area of each strip is equal to the height f(X), times the width dX. Hence we are taking the sum (integral) of all the strips of area f(X)dX, as x goes from a to b.

<u>Integration and differentiation</u>

Why is finding the area under a curve the reverse of differentiation?

Well, let f(X) be our function that we're looking for the area under, and let F(x) be the area under the curve between X=a and X=x.

What is dF/dx? This is the *rate of change* of F(x) as x increases. In other words, it is the *marginal area* as x increases. But this is simply equal to the height of the graph at this point, in other words, it is equal to f(x). Hence dF/dx =f(x), so the area under the curve f(x) is found by finding a function F(x) whose differential is f(x).

Specifically, let f(x) be a (continuous) function. Let F(x) be a function such that dF/dx=f(x).

Then the area under the curve of f(x) between a and b,

$$\int_{X=a}^{X=b} f(X)dX = F(b) - F(a)$$

<u>Example</u>

Find $\int_{x=1}^{x=4} x^2 dx$, the area under the curve of x^2 between x=1 and x=4.

Let $f(x)=x^2$

The (indefinite) integral of f(x) is $F(x)=x^3/3$.

Hence the area under the curve (the definite integral) is equal to F(4)-F(1) = $4^3/3 - 1^3/3 = 64/2 - 1/3 = 21$.

<u>Double integrals</u>

Frequently in statistics we may need to integrate functions of two variables, for example in calculating probabilities from joint probability distributions.

A <u>double integral</u> is essentially the integral of an integral.

That is, if f(X,Y) is a function of two variables, then we can calculate

$\int_{X=a}^{X=b} \int_{Y=c}^{Y=d} f(X,Y)dYdX$ - first we perform the inner integral, that is we integrate the function with respect to Y, to get a function of X; then we integrate the result with respect to X.

This can be seen as finding the <u>volume</u> under the surface given by f(X,Y), between the bounds x=a to x=b, and y=c to y=d.

Example

Let f(X,Y)=XY. Find $\int_{X=0}^{X=1}\int_{Y=0}^{Y=1} f(X,Y)dYdX$

First we need to integrate w.r.t. Y, that is, we need to find a function F(X,Y) s.t. $\frac{\partial F}{\partial Y}$=XY

We may take F(X,Y)=$XY^2/2$. Then $\int_{Y=0}^{Y=1} XYdY$ = F(X,1)-F(X,0) = $(X*1^2)/2 - (X*0^2)/2$ = X/2.

So we can now place this answer, X/2, into the outer integral, and find:

$$\int_{X=0}^{X=1} \frac{X}{2}dX$$

Let g(x) = X/2. Then the integral is G(X)=$X^2/4$, so that dG/dX=g(x).
Hence the integral is equal to G(1)-G(0) = $1^2/4$ = ¼.