INTRODUCTION TO MATHEMATICAL STATISTICS II  
Semester 1 2001  

Some notes on multi-variable integrals

Recall from Maths 1 the definition of the integral of a continuous function $f: [a, b] \to \mathbb{R}$. We divide the interval $[a, b]$ into $n$ subintervals $[a = t_0, t_1], [t_1, t_2], \ldots, [t_{n-1}, t_n = b]$ of equal size. Notice that the length of the interval $[t_i, t_{i+1}]$ is $t_{i+1} - t_i = (b-a)/n$. We form the upper sum $U_n(f)$ as

$$U_n(f) = \sum_{i=1}^{n} \max\{f(x) \mid x \in [t_{i-1}, t_i]\} \frac{b-a}{n}$$

and the lower sum

$$L_n(f) = \sum_{i=1}^{n} \min\{f(x) \mid x \in [t_{i-1}, t_i]\} \frac{b-a}{n}$$

It is a theorem that there exists a unique number $L$ such that for every $n$, $L_n(f) \leq L \leq U_n(f)$ and that as $n \to \infty$ $L_n(f)$ gets closer and closer to $L$ from below and $U_n(f)$ gets closer and closer to $L$ from above. We call $L$ the integral of $f$ over $[a, b]$ and write

$$L = \int_{a}^{b} f(x)dx.$$ 

Consider now a closed, bounded region $R \subset \mathbb{R}^2$. Because $R$ is bounded it is a subset of a rectangle $[a, b] \times [c, d]$. For each $n$ we divide $[a, b]$ into $n$ intervals $[a = x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n = b]$ and similarly we divide $[c, d]$ into $n$ intervals $[c = y_0, y_1], [y_1, y_2], \ldots, [y_{n-1}, y_n = d]$. This divides the rectangle $[a, b] \times [c, d]$ into $n^2$ rectangles of area $(b-a)(d-c)/n^2$. Let

$$R_{ij} = R \cap [x_{i-1}, x_i] \times [y_{j-1}, y_j].$$

Notice that if $R$ is an irregular shape it may be that some of the $R_{ij}$ are empty sets, that is $R_{ij} = \emptyset$.

We then define

$$U_n(f) = \sum_{i,j=1, R_{ij} \neq \emptyset}^{n} \max\{f(x, y) \mid (x, y) \in R_{ij}\} \frac{(b-a)(d-c)}{n^2}$$

and the lower sum

$$L_n(f) = \sum_{i,j=1, R_{ij} \neq \emptyset}^{n} \min\{f(x, y) \mid (x, y) \in R_{ij}\} \frac{(b-a)(d-c)}{n^2}$$

Just as in the case of one variable if $f$ is continuous we find that there is a unique $L$ such that for every $n$, $L_n(f) \leq L \leq U_n(f)$ and that as $n \to \infty$ $L_n(f)$ gets closer and closer to $L$ from below and $U_n(f)$ gets closer and closer to $L$ from above. We call $L$ the integral of $f$ over $[a, b]$ and write

$$L = \int_{R} f(x)dA.$$ 

Just as in the one-variable case the definition gives you an idea of how to calculate the integral numerically but is of little use if you want a closed form answer. In the case of one variable we use the anti-derivatives and the Fundamental Theorem of
Calculus. In the case of two (or more) variables we reduce to the one-variable case as follows.

For convenience let us assume that $R = [a, b] \times [c, d]$. We expect the integral to be approximately equal to

$$\sum_{i,j=1}^{n} f(x_{i}^{*}, y_{j}^{*}) \frac{(b-a)(d-c)}{n^2}$$

where $(x_{i}^{*}, y_{j}^{*})$ is any point in $R_{ij}$. This sum can be performed by first summing over $i$ and then $j$. This looks like

$$\sum_{j=1}^{n} \left( \sum_{i=1}^{n} f(x_{i}^{*}, y_{j}^{*}) \frac{(b-a)}{n} \right) \frac{(c-d)}{n}$$

Notice that the inner sum looks like it is approximately $\int_{a}^{b} f(x, y_{j}^{*}) dx$. The second sum is then approximately the integral over $[c, d]$ of $\int_{a}^{b} f(x, y) dx$ thought of as a function of $y$. It is in fact possible to prove that if $f$ is continuous on a closed and bounded region $R$ then

$$\int \left( \int f(x, y) dx \right) dy = \int_{R} f(x, y) dA = \int \left( \int f(x, y) dy \right) dx.$$ 

This is how we calculate $\int_{R} f(x, y) dA$: we integrate the $x$ and the $y$ variables one after the other in either order. Notice that we have omitted the limits of the one variable integrals because they will typically be functions of the other variable. That is, for given $y$ the set of all $(x, y)$ which lie in $R$ will be of the form $a(y) \leq x \leq b(y)$ and similarly for $x$. 