Integration

- Univariate integrals. The R command for computing $\int_a^b f(x)dx$ is

```
integrate(f,a,b)$value
```

Here a can be $-\infty$ and b can be ∞:

```
integrate(f,-Inf,b)$value    #a=-Inf
integrate(f,a,Inf)$value     #b=Inf
```

Example 1. Let

$$f(x) = x^2\sqrt{1-x^2} + \frac{x(1-x^2)}{2}$$

for $x \in [0,1]$. Write down the R commands for computing $\int_0^1 f(x)dx$.

Solution.

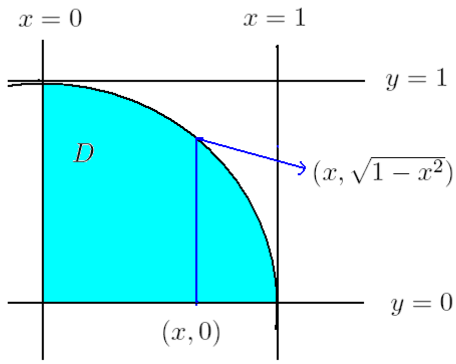
```
f <- function(x){ return(x^2*sqrt(1-x^2)+x*(1-x^2)/2) }
integrate(f,0,1)$value #pi/16+1/8
```

- Bivariate integrals.
  - To compute $\int_D f(x,y)d(x,y)$, where D is a bounded region, we can define
    $$g(x,y) = \begin{cases} f(x,y) & \text{if } (x,y) \in D; \\ 0 & \text{otherwise,} \end{cases}$$
    and find intervals $[a,b]$ and $[c,d]$ such that $D \subset [a,b] \times [c,d]$, then
    $$\int_D f(x,y)d(x,y) = \int_a^b \int_c^d g(x,y)dydx.$$
    To find the integral, we need to first define
    $$h(x) = \int_c^d g(x,y)dy \tag{1}$$
    and then compute $\int_a^b h(x)dx$.
  - Note: In (1), sometimes
    $$h(x) = \int_c^d g(x,y)dy = \int_{\ell(x)}^{u(x)} f(x,y)dy$$
    and h can be computed using a more precise integration range.

- Example 2. Let

$$D = \{(x, y) : x^2 + y^2 \le 1, x \ge 0, y \ge 0\}.$$

Find $\int_D (x^2 + xy) d(x, y)$. The graph of D is shown below.



– Solution 1. Let

$$g(x, y) = \begin{cases} x^2 + xy & \text{if } (x, y) \in D; \\ 0 & \text{otherwise.} \end{cases}$$

Since $D \subset [0, 1] \times [0, 1]$, we have

$$\int_D (x^2 + xy) d(x, y) = \int_0^1 \underbrace{\int_0^1 g(x, y) dy}_{h(x)} dx. \qquad (2)$$

R commands for computing $h(x) = \int_0^1 g(x, y) dy$ and $\int_0^1 h(x) dx$:

```
g <- function(x,y){
  ans <- x^2+x*y
  ans[x<0] <- 0
  ans[y<0] <- 0
  ans[(x^2+y^2)>1] <- 0
  return(ans)
}

h0 <- function(x){
   gx <- function(y){ return(g(x,y)) }
   return( integrate(gx, 0, 1)$value )
 }
```

```
h <- function(x){
 n <- length(x)
 ans <- rep(0, n)
 for (i in 1:n){ ans[i] <- h0(x[i]) }
 return(ans)
}
#or h <- Vectorize(h0)

integrate(h, 0, 1)$value
pi/16 + 1/8
#the true value is pi/16 + 1/8
```

– Solution 2. In (2),

$$h(x) = \int_0^1 g(x,y)dy = \int_0^{\sqrt{1-x^2}} (x^2 + xy)dy,$$

so the R commands for computing $h(x)$ and $\int_0^1 h(x)dx$ can be modified accordly:

```
g <- function(x,y){
  ans <- x^2+x*y
  ans[x<0] <- 0
  ans[y<0] <- 0
  ans[(x^2+y^2)>1] <- 0
  return(ans)
}

h0 <- function(x){
   gx <- function(y){ return(x^2+x*y) }
   return( integrate(gx, 0, sqrt(1-x^2))$value )
 }
h <- Vectorize(h0)
integrate(h, 0, 1)$value
pi/16 + 1/8
```

– Note. In Solution 2, the function h is the function f in Example 1.

- 多重積分 (multivariate integrals). 在函數爲可積的條件下,

$$\begin{aligned} &\int_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]} f(x_1,x_2,\ldots,x_k)d(x_1,x_2,\ldots,x_k) \\ &= \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_k}^{b_k} f(x_1,x_2,\ldots,x_k)dx_k\cdots dx_2dx_1. \end{aligned}$$

但若直接依此式進行計算非常耗費計算資源。

- Use the Monte Carlo method (蒙地卡羅方法) to compute

$$I = \int_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]} f(x_1,x_2,\ldots,x_k)d(x_1,x_2,\ldots,x_k). \quad (3)$$

- Idea of the Monte Carlo method for evaluating $\int g(x)dx$.

  Find a density f_0 and generate IID data X_1, …, X_n so that each X_i has PDF f_0, then

$$\int g(x)dx = \int \frac{g(x)}{f_0(x)} f_0(x)dx = E\left(\frac{g(X_1)}{f_0(X_1)}\right) \approx \frac{1}{n}\sum_{i=1}^{n} \frac{g(X_i)}{f_0(X_i)}$$

  for large n. Here n is called the Monte Carlo sample size, g is a function on R^k and $\int g(x)dx$ means $\int_{R^k} g(x)dx$. Note that $g(x)/f_0(x)$ can be defined to be any function of x when $f_0(x) = 0$.

- To compute I in (3), note that $I = \int g(x)dx$ with

$$g(x_1,\ldots,x_k) = I_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]}(x_1,\ldots,x_k)f(x_1,\ldots,x_k),$$

$$I_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]}(x_1,\ldots,x_k) = \begin{cases} 1 & \text{if } x_j \in [a_j,b_j] \text{ for } j=1,\ldots,k; \\ 0 & \text{otherwise.} \end{cases}$$

  Let U_1, …, U_k be independent random variables such that $U_j \sim U(a_j,b_j)$ for $j = 1$, …, k, and let

$$f_0(x_1,\ldots,x_k) = I_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]}(x_1,\ldots,x_k)\frac{1}{(b_1-a_1)\cdots(b_k-a_k)}$$

  then f_0 is the joint density of U_1, …, U_k,

$$\frac{g(x)}{f_0(x)} = (b_1-a_1)\cdots(b_k-a_k)f(x)$$

  and

$$\begin{aligned} I &= \int_{[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_k,b_k]} f(x_1,x_2,\ldots,x_k)d(x_1,x_2,\ldots,x_k) \\ &= \int g(x)dx \\ &\approx \frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i)}{f_0(X_i)} \\ &= (b_1-a_1)\cdots(b_k-a_k)\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right) \end{aligned}$$

  for large n, where X_1, …, X_n are IID and each $X_i \sim (U_1,\ldots,U_k)$ has PDF f_0.

- Example 3. Let

$$D = \{(x, y) : x^2 + y^2 \le 1, x \ge 0, y \ge 0\}.$$

Use Monte Carlo method to compute

$$\int_D (x^2 + xy)d(x, y) = \int_{[0,1]\times[0,1]} g(x, y)d(x, y),$$

where

$$g(x,y) = \begin{cases} x^2 + xy & \text{if } (x,y) \in D; \\ 0 & \text{otherwise.} \end{cases}$$

Take the Monte Carlo sample size to be 10^6.

Sol. Consider generate Monte Carlo sample from the distribution of (U_1, U_2), where U_1 and U_2 are independent $U(0, 1)$ random variables.

```
g <- function(x,y){
  ans <- x^2+x*y
  ans[x<0] <- 0
  ans[y<0] <- 0
  ans[(x^2+y^2)>1] <- 0
return(ans)
   }
x <- runif(10^6, 0, 1)
y <- runif(10^6, 0, 1)
mean(g(x,y))*(1-0)*(1-0)
pi/16+1/8
```

- Practice problems.

1. (10 pts) Let

$$D = \{(x, y) : x \ge 0, y \ge 0, x + y \le 2\}.$$

Write down the R commands for computing the integral

$$\int_D \sin(x^2 + y^2)d(x, y).$$

You may use Monte Carlo method to find the integral or compute the integral as $\int_0^2 h(x)dx$ for some function h. Take the Monte Carlo sample size to be 10^6 if you use Monte Carlo method to find the integral.

2. (10 pts) Let $D = \{(x, y, z) : x^2 + y^2 + z^2 \le 1, x \ge 0, y \ge 0, z \ge 0\}$ and

$$g(x,y,z) = \begin{cases} 1 & \text{if } (x,y,z) \in D; \\ 0 & \text{otherwise.} \end{cases}$$

Write down the R commands for computing the integral

$$\int_{[0,1]\times[0,1]\times[0,1]} g(x,y,z)d(x,y,z)$$

using Monte Carlo method. Take the Monte Carlo sample size to be 10^6.