U-substitution, integration by parts and numerical integration Lecture 4 – 2021-05-14

MAT A35 – Summer 2021 – UTSC

Prof. Yun William Yu

Substitution rule algorithm

$\sin^2 x = (\sin x)^2$

- $\boxed{\text{Step 1: Guess an appropriate } u}$
- Step 2: Compute du, dx, and x
- Step 3: Substitute in to get rid of all the x's
- Step 4: Integrate as a function of u
- Step 5: Convert back to x's

What is a useful change of variables?

Ex. $\int \frac{\boxed{3x^2 dx}}{1+x^3}$

$u = 1 + x^3$

$du = \boxed{3x^2 dx}$

$\int \frac{du}{u}$

$u = 3x^2$

$du = 6x\,dx$

Ex. $\int \sin^2 x \boxed{\cos x\, dx}$

$u = \sin x$

$du = \boxed{\cos x\, dx}$

$\int u^2 du$

$u = \cos x \Rightarrow x = \cos^{-1} u$

$du = -\sin x\, dx$

$-\int \sin x \cos x\, du$

$= -\int [\sin(\cos^{-1} u)] u\, du$

Want a change of variables $u = u(x)$ such that after substituting in u and du, the equation is simpler. Often, a good sign is if u and du appear in the original integral, up to a constant.

Try it out: what's the u-substitution?

- $\int x^4 e^{x^5} dx$

$u = x^5$

$du = 5x^4 dx$

$= \int \frac{1}{5} e^u \, du$

$= \frac{1}{5} e^u + C = \frac{1}{5} e^{x^5} + C$

A: $u = x^4$
B: $u = x^5$
C: $u = e^{x^5}$
D: $u = x^4 dx$
E: None

- $\int e^{4x} dx$

$u = 4x$

$du = 4dx$

$= \int \frac{1}{4} e^u du$

$= \frac{1}{4} e^u + C = \frac{1}{4} e^{4x} + C$

A: $u = 4x$
B: $u = e^{4x}$
C: $u = e^x$
D: $u = e^{4x} dx$
E: None

- $\int 12x\sqrt{1 + 6x^2} dx$

$u = 1 + 6x^2$

$u = 6x^2$

$du = 12x dx$

$= \int \sqrt{1+u} \, du$

$v = 1 + u$

$dv = du$

$= \int v^{\frac{1}{2}} dv$

A: $u = 12x$
B: $u = 1 + 6x^2$
C: $u = \sqrt{1 + 6x^2}$
D: $u = 6x^2$
E: None

Integration by parts algorithm

- $\int u\, dv = uv - \int v\, du$
- Step 1: Guess which part is u and which part is dv
- Step 2: Apply the formula above and hope you can solve $\int v\, du$
- Step 3: If it doesn't, try again with a different guess for u and dv.

- Step ?: Give up if no guess seems to work. The integral might not be amenable to integration by parts.

$$\int u\,dv = uv - \int v\,du$$

$$\int \ln x\,dx = x \ln x - \int dx = x \ln x - x + C$$

$u = \ln x$ $\qquad$ $v = \int dx = x$

$du = \frac{1}{x}dx$ $\qquad$ $dv = dx$

$$\int x \cos x\,dx = x \sin x - \int \sin x\,dx = x \sin x + \cos x + C$$

$u = x$ $\qquad$ $v = \int \cos x\,dx = \sin x$

$du = 1\,dx$ $\qquad$ $dv = \cos x\,dx$

Integration by parts heuristic: DETAIL

Functions near the top of the list have easy antiderivatives, so are good guesses for dv.

- D: (dv)
- E: exponential functions ($e^{2x}, 2^x$)
- T: trigonometric functions ($\sin x$, $\tan x$, $\operatorname{sech} x$)
- A: algebraic functions (x^2, $2(x+1)^2$)
- I: inverse trigonometric functions ($\arcsin x$, $\operatorname{arccosh} x$)
- L: logarithmic functions ($\ln x$, $\log_{10} 2x$)

Functions near the bottom of the list have easy derivatives, so are good guesses for u.

NB: there are many exceptions to this heuristic. (e.g. sometimes I and L are swapped, and sometimes you need to split algebraic functions into two pieces)

DETAIL example ($\int u\,dv = uv - \int v\,du$)

$$\int x^2 e^{-2x}\,dx = x^2\cdot\left(-\frac{1}{2}e^{-2x}\right) - \int\left(-\frac{1}{2}e^{-2x}\right)2x\,dx$$

$u = x^2$ $\quad v = -\frac{1}{2}e^{-2x}$

$du = 2x\,dx$ $\quad dv = e^{-2x}\,dx$

$$= -\frac{x^2}{2}e^{-2x} + \int x e^{-2x}\,dx$$

$u = x$ $\quad v = -\frac{1}{2}e^{-2x}$

$du = dx$ $\quad dv = e^{-2x}\,dx$

$$= -\frac{x^2}{2}e^{-2x} + \left[-\frac{x}{2}e^{-2x} - \int\left(-\frac{1}{2}e^{-2x}\right)dx\right]$$

$$= -\frac{x^2}{2}e^{-2x} - \frac{x}{2}e^{-2x} + \frac{1}{2}\int e^{-2x}\,dx$$

$$= -\frac{x^2}{2}e^{-2x} - \frac{x}{2}e^{-2x} - \frac{1}{4}e^{-2x} + C$$

Try it out: $\int_1^e x \ln x^2 \, dx$

$\ln x^2 = 2 \ln x$

- Hints: $\int u \, dv = uv - \int v \, du$ or $\int_{x=a}^{x=b} u \, dv = uv|_{x=a}^{x=b} - \int_{x=a}^{x=b} v \, du$
- DETAIL (dv, exp, trig, algebraic, inverse trig, log)

$u = \ln x^2$ $\qquad v = \frac{1}{2}x^2$

$du = \frac{1}{x^2} \cdot 2x = \frac{2}{x}$ $\qquad dv = x\,dx$

$\int_1^e x \ln x^2 \, dx = \frac{1}{2}x^2 \ln x^2 \Big|_1^e - \int_1^e x \, dx = \frac{1}{2}x^2 \ln x^2 - \frac{1}{2}x^2 \Big|_1^e$

$= \left[\frac{1}{2}e^2 \underbrace{\ln e^2}_{2} - \frac{1}{2}e^2\right] - \left[-\frac{1}{2}\right]$

$= \frac{1}{2}e^2 + \frac{1}{2}$

A: $e^2 + 1$

B: $\frac{e^2+1}{2}$

C: $e^2 - 1$

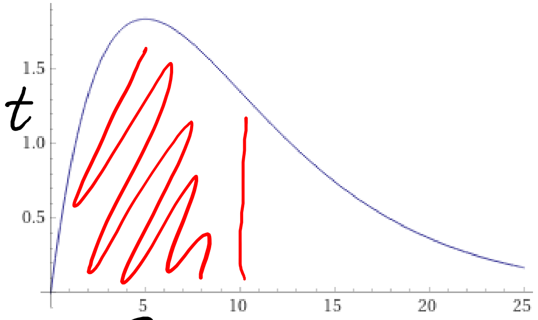
D: $\frac{e^2-1}{2}$

E: None

Application – Drug dosage

- Suppose a patient takes 25mg of a drug orally and it is metabolized from the body at a rate of $E(t) = te^{-kt}$, where $k = 0.2$ mg/hour and t is time in hours since taking the drug.
- How much drug has been metabolized after 10 hours?

$$\int_0^{10} t e^{-kt}\,dt = -\frac{t}{k}e^{-kt}\Big|_0^{10} - \int_0^{10} -\frac{1}{k}e^{-kt}\,dt$$



$u = t \qquad v = -\frac{1}{k}e^{-kt}$

$du = dt \qquad dv = e^{-kt}\,dt$

$$= -\frac{10}{0.2}e^{-2} - \left[\frac{1}{k^2}e^{-kt}\right]\Big|_0^{10}$$

$$= -50e^{-2} - \left[25e^{-2} - 25\right]$$

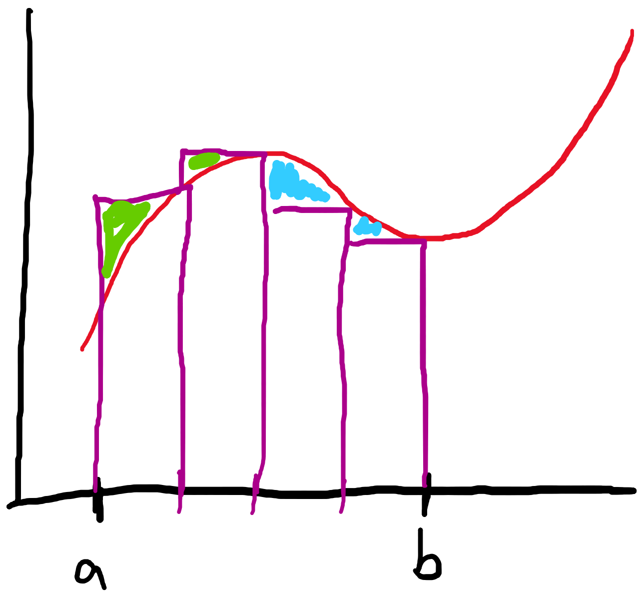
$$= 25 - 75e^{-2} \approx 14.85 \text{ mg.}$$

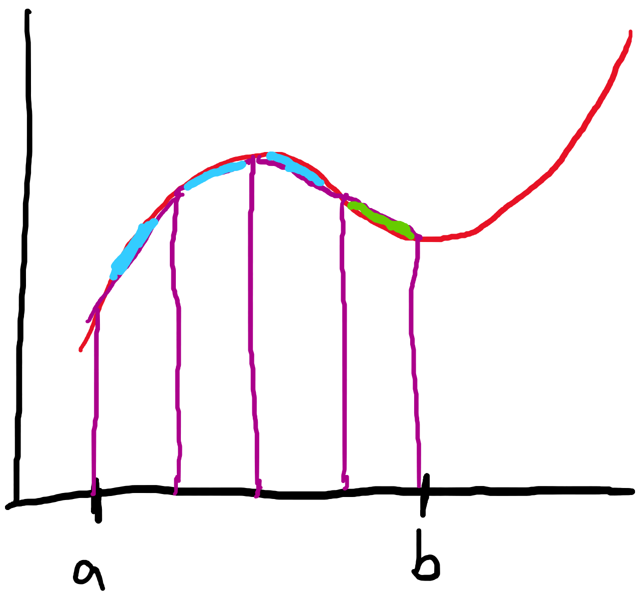
Theory vs practice

- Practical tools
  - Integral tables (need change of variables/u-substitution)
    - Table 1, pg. 748, in textbook (Bittinger, Brand, Quintanilla)
    - http://integral-table.com/downloads/single-page-integral-table.pdf
  - Calculators:
    - Desmos: https://www.desmos.com/calculator/be5ne9vwi8
    - WolframAlpha: https://www.wolframalpha.com/input/?i=what+is+the+integral+of+%28x%2B1%29%5E2+ln+%28x%2B1%29
- Why should you practice what a calculator can do?
  - Building blocks for more advanced techniques/analyses.
  - Intuition for when things go wrong.
  - Understanding how the calculators work so you can modify the algorithm when faced with a (slightly) different problem.

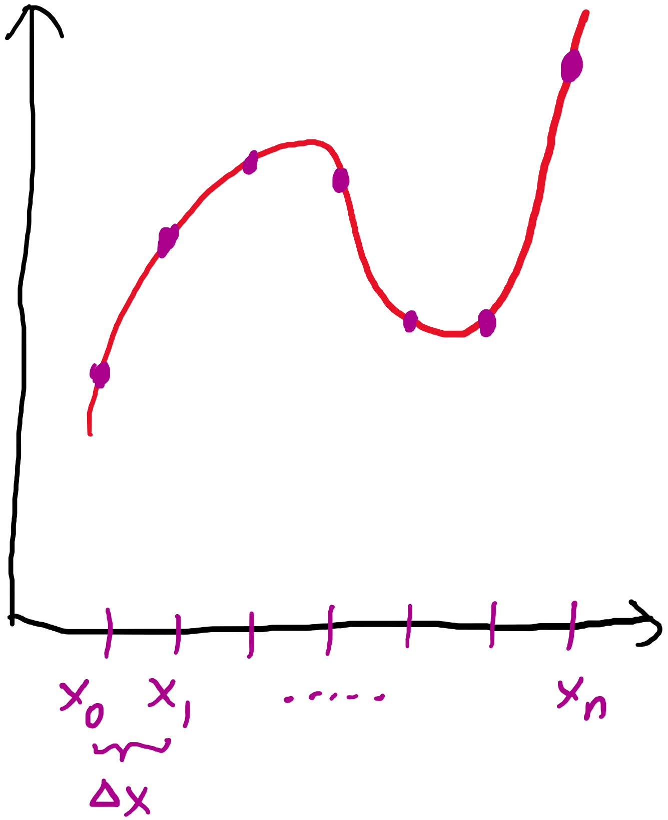
Numerical integration

- We can approximate area under any curve by dividing into shapes we know how to compute area for, like rectangles or trapezoids

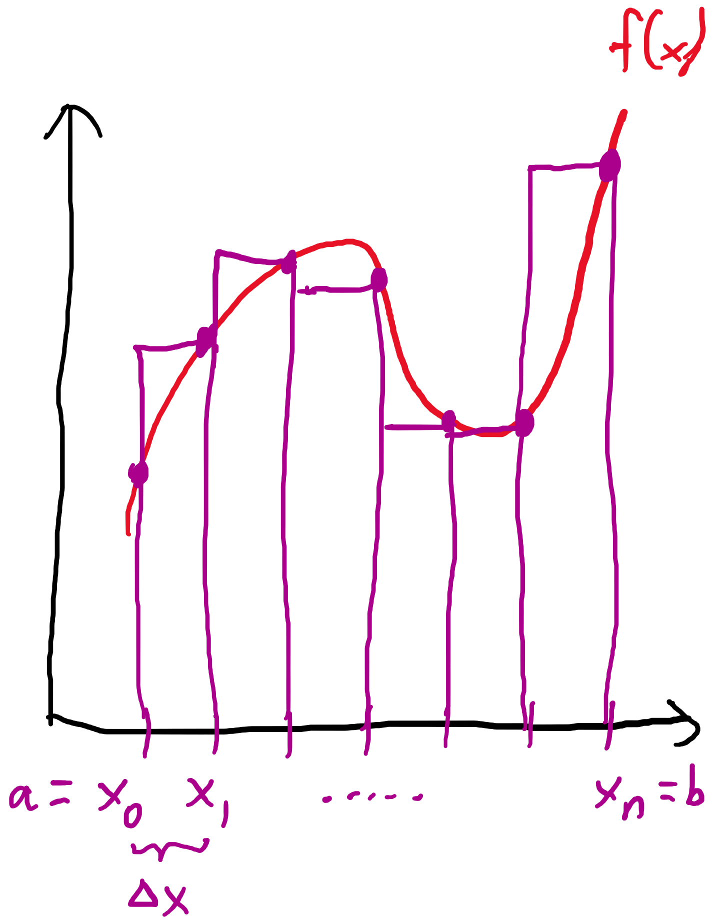




Riemann summation rule



Riemann summation rule

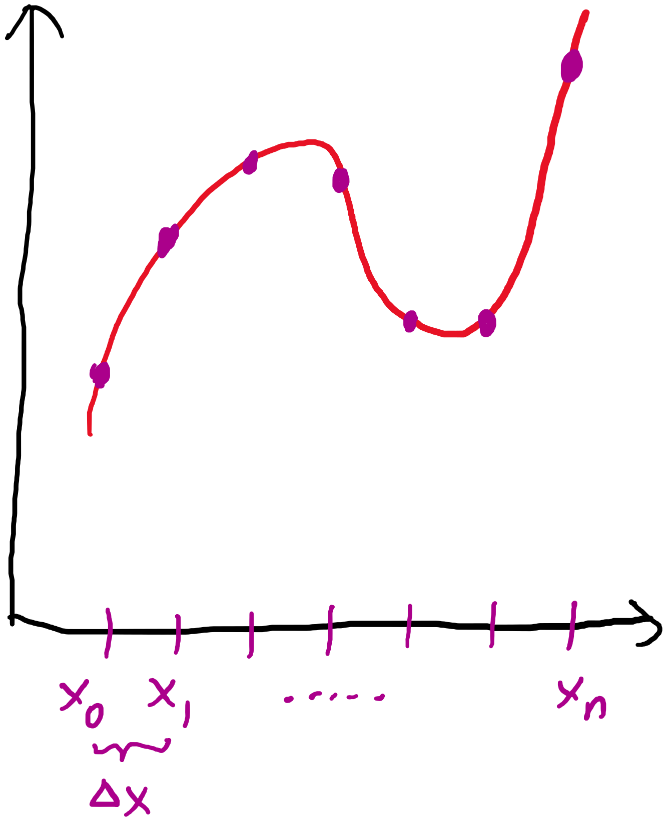


$$\text{Area} = \sum_{i=1}^{n} \Delta x \, f(x_i) = \Delta x \sum_{i=1}^{n} f(x_i)$$

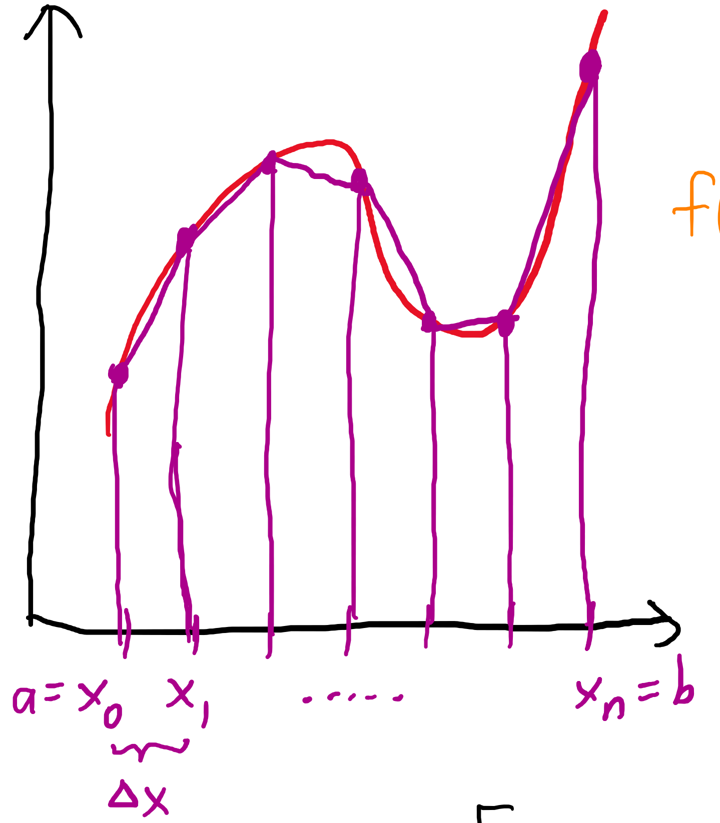
i	$x_i = a + i\Delta x$	$f(x_i)$	Weight	Term
0	x_0	$\vdots$	0	weight × $f(x_i)$
1	x_1	$\vdots$	1	$\vdots$
2	x_2	$\vdots$	1	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	x_n		1	

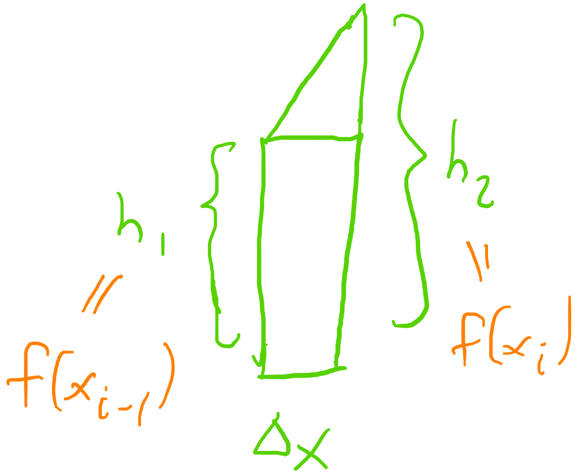
$$\sum_{i=1}^{n} f(x_i)$$

Trapezoid rule



Trapezoid rule





Area =

$$\Delta x \cdot h_1 + \frac{1}{2}\Delta x (h_2 - h_1) = \frac{1}{2}\Delta x (h_1 + h_2)$$

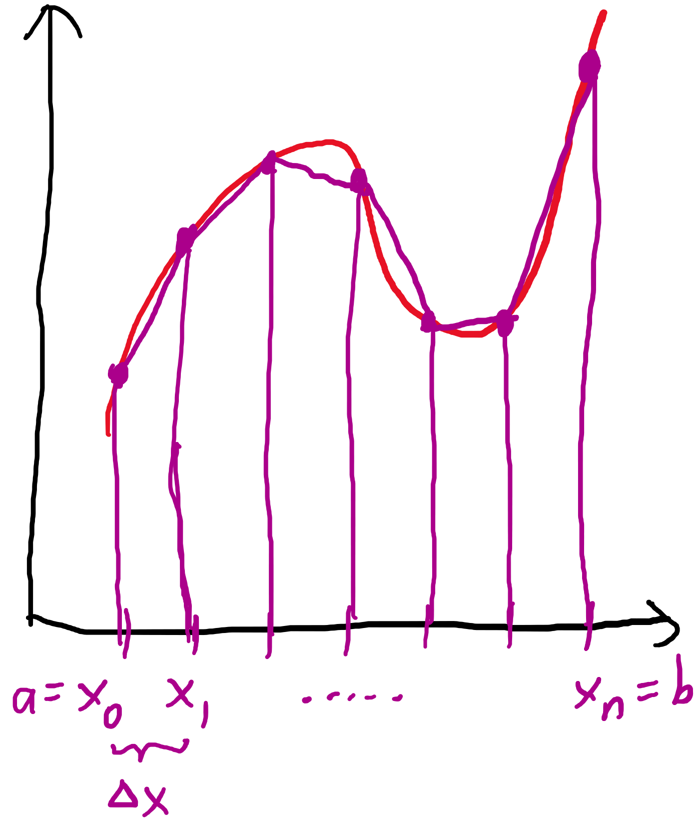
Area under curve $\approx$

$$\sum_{i=1}^{n} \frac{1}{2}\Delta x \left(f(x_i) + f(x_{i-1}) \right)$$

$$= \frac{1}{2}\Delta x \left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n) \right]$$

Trapezoid rule

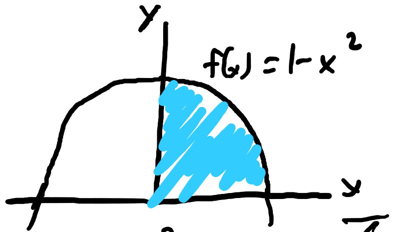
$$\frac{1}{2}\Delta x\left[f(x_0)+2f(x_1)+2f(x_2)+\dots+2f(x_{n-1})+f(x_n)\right]$$



i	x_i	$f(x_i)$	Weight	Term
0	x_0		1	Weight × $f(x_i)$
1	x_1		2	
2	x_2		2	
⋮	⋮		⋮	
$n-1$	x_{n-1}		2	
n	x_n		1	

$[\dots]$

Example: $\int_0^1 (1 - x^2)dx\,, n = 10$



AUC = $\frac{2}{3} = 0.66\bar{6}$

- $a = 0, b = 1, \Delta x = 0.1$, and $f(x) = 1 - x^2$

i	$x_i = a + i\Delta x$	$f(x_i)$	Riemann weight	Riemann Term	Trapezoid weight	Trapezoid Term
0	0	1	0	1	1	1
1	0.1	0.99	1	0.99	2	1.98
2	0.2	0.96	1	0.96	2	1.92
3	0.3	0.91	1	0.91	2	1.82
4	0.4	0.86	1	0.86	2	1.72
5	0.5	0.75	1	0.75	2	1.50
6	0.6	0.64	1	0.64	2	1.28
7	0.7	0.51	1	0.51	2	1.02
8	0.8	0.36	1	0.36	2	0.72
9	0.9	0.19	1	0.19	2	0.38
10	1	0	1	0	1	0

Sum: 6.17 Sum: 13.34

Riemann area: 0.1×6.17
$= 0.617$

Trap area: $0.1 \times \frac{1}{2} \times 13.34$
$= 0.667$

Simpson's Rule of Thirds (parabolic)

- $\int_0^1 (1 - x^2)dx\,, n = 10\,, a = 0, b = 1, \Delta x = 0.1$, and $f(x) = 1 - x^2$

i	$x_i = a + i\Delta x$	$f(x_i)$	Riemann weight	Riemann Term	Trapezoid weight	Trapezoid Term	Simpson weight	Simpson Term
0	0	1	0	0	1	1	1	1
1	0.1	0.99	1	0.99	2	1.98	4	3.96
2	0.2	0.96	1	0.96	2	1.92	2	1.92
3	0.3	0.91	1	0.91	2	1.82	4	3.64
4	0.4	0.86	1	0.86	2	1.72	2	1.72
5	0.5	0.75	1	0.75	2	1.50	4	3.00
6	0.6	0.64	1	0.64	2	1.28	2	1.28
7	0.7	0.51	1	0.51	2	1.02	4	2.04
8	0.8	0.36	1	0.36	2	0.72	2	0.72
9	0.9	0.19	1	0.19	2	0.38	4	0.76
10	1	0	1	0	1	0	1	0

Riemann area: 0.617

Trap area: 0.667

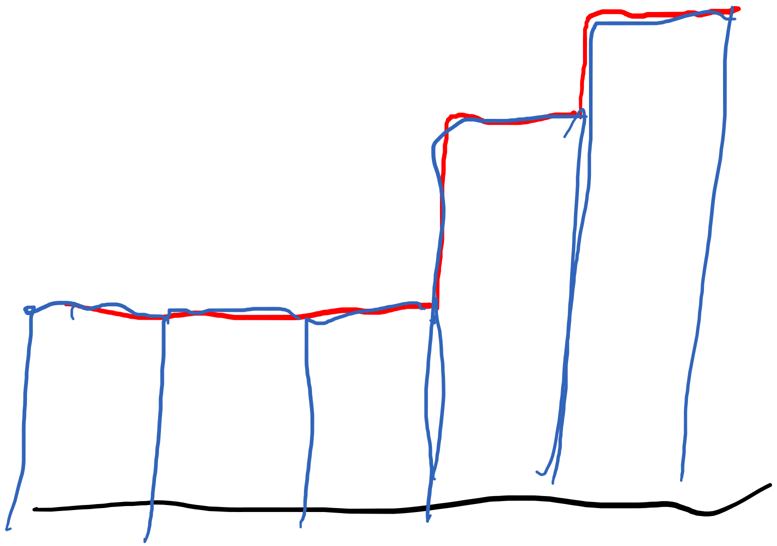
Simpson Sum: 20.04

Simpson Area: $\frac{1}{3} \cdot 0.1 \cdot 20.04$
$= 0.668$

Most accurate approximation

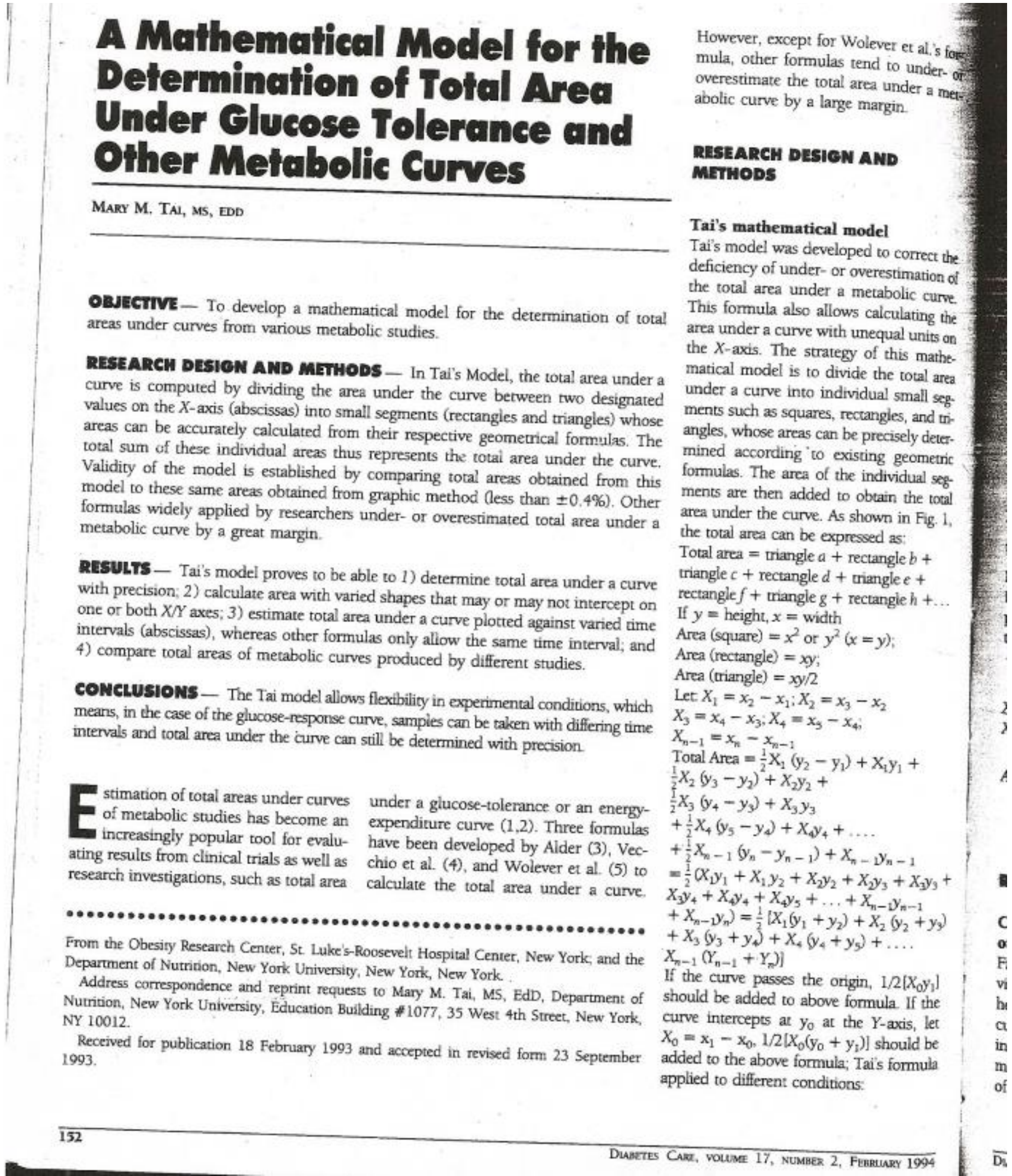
- Which approximation is most accurate?

A: Riemann
B: Trapezoid
C: Simpson
D: B or C
E: None



- The accuracy of an approximation depends on the function being approximated.

Area under experimentally sampled curve



A Mathematical Model for the Determination of Total Area Under Glucose Tolerance and Other Metabolic Curves

MARY M. TAI, MS, EDD

OBJECTIVE — To develop a mathematical model for the determination of total areas under curves from various metabolic studies.

RESEARCH DESIGN AND METHODS — In Tai's Model, the total area under a curve is computed by dividing the area under the curve between two designated values on the X-axis (abscissas) into small segments (rectangles and triangles) whose areas can be accurately calculated from their respective geometrical formulas. The total sum of these individual areas thus represents the total area under the curve. Validity of the model is established by comparing total areas obtained from this model to these same areas obtained from graphic method (less than ±0.4%). Other formulas widely applied by researchers under- or overestimated total area under a metabolic curve by a great margin.

RESULTS — Tai's model proves to be able to *1*) determine total area under a curve with precision; *2*) calculate area with varied shapes that may or may not intercept on one or both X/Y axes; *3*) estimate total area under a curve plotted against varied time intervals (abscissas), whereas other formulas only allow the same time interval; and *4*) compare total areas of metabolic curves produced by different studies.

CONCLUSIONS — The Tai model allows flexibility in experimental conditions, which means, in the case of the glucose-response curve, samples can be taken with differing time intervals and total area under the curve can still be determined with precision.

Estimation of total areas under curves of metabolic studies has become an increasingly popular tool for evaluating results from clinical trials as well as research investigations, such as total area under a glucose-tolerance or an energy-expenditure curve (1,2). Three formulas have been developed by Alder (3), Vecchio et al. (4), and Wolever et al. (5) to calculate the total area under a curve.

From the Obesity Research Center, St. Luke's-Roosevelt Hospital Center, New York; and the Department of Nutrition, New York University, New York, New York.

Address correspondence and reprint requests to Mary M. Tai, MS, EdD, Department of Nutrition, New York University, Education Building #1077, 35 West 4th Street, New York, NY 10012.

Received for publication 18 February 1993 and accepted in revised form 23 September 1993.

However, except for Wolever et al.'s formula, other formulas tend to under- or overestimate the total area under a metabolic curve by a large margin.

RESEARCH DESIGN AND METHODS

Tai's mathematical model

Tai's model was developed to correct the deficiency of under- or overestimation of the total area under a metabolic curve. This formula also allows calculating the area under a curve with unequal units on the X-axis. The strategy of this mathematical model is to divide the total area under a curve into individual small segments such as squares, rectangles, and triangles, whose areas can be precisely determined according to existing geometric formulas. The area of the individual segments are then added to obtain the total area under the curve. As shown in Fig. 1, the total area can be expressed as:

Total area = triangle a + rectangle b + triangle c + rectangle d + triangle e + rectangle f + triangle g + rectangle h +…

If y = height, x = width

Area (square) = x^2 or y^2 $(x = y)$;

Area (rectangle) = xy;

Area (triangle) = $xy/2$

Let: $X_1 = x_2 - x_1$; $X_2 = x_3 - x_2$

$X_3 = x_4 - x_3$; $X_4 = x_5 - x_4$;

$X_{n-1} = x_n - x_{n-1}$

$$\begin{aligned}\text{Total Area} &= \tfrac{1}{2}X_1(y_2 - y_1) + X_1y_1 + \tfrac{1}{2}X_2(y_3 - y_2) + X_2y_2 + \tfrac{1}{2}X_3(y_4 - y_3) + X_3y_3 \\ &+ \tfrac{1}{2}X_4(y_5 - y_4) + X_4y_4 + \dots + \tfrac{1}{2}X_{n-1}(y_n - y_{n-1}) + X_{n-1}y_{n-1} \\ &= \tfrac{1}{2}(X_1y_1 + X_1y_2 + X_2y_2 + X_2y_3 + X_3y_3 + X_3y_4 + X_4y_4 + X_4y_5 + \dots + X_{n-1}y_{n-1} + X_{n-1}y_n) \\ &= \tfrac{1}{2}[X_1(y_1 + y_2) + X_2(y_2 + y_3) + X_3(y_3 + y_4) + X_4(y_4 + y_5) + \dots X_{n-1}(Y_{n-1} + Y_n)]\end{aligned}$$

If the curve passes the origin, $1/2[X_0y_1]$ should be added to above formula. If the curve intercepts at y_0 at the Y-axis, let $X_0 = x_1 - x_0$, $1/2[X_0(y_0 + y_1)]$ should be added to the above formula; Tai's formula applied to different conditions:

- What if we don't have an exact formula for a curve, but just samples along it?
- We can still treat our discrete measurements as samples of $f(x_i)$.
- i.e. even when explicit integration fails, understanding the ideas behind integration lets you apply the related approximations.