Example In the next assignment, we may want to find the expected price of oil in December, given an assumption that we know its probability density function \( p(x) \). That's

\[
E[\text{price}] = \int_{-\infty}^{\infty} x \cdot p(x) \, dx
\]

or, say, the probability that the price exceeds a given value \( V \)

\[
\text{prob. \{price exceeds } V \} = \int_{V}^{\infty} p(x) \, dx
\]

If \( p(x) \) is something we can't integrate analytically (Freshman calculus), then we must resort to numerical methods.

Problem goes back 200 years to the masters, Newton, Gauss, Euler!

References used

[Acton 96] "Real Computing Made Real"

- lots of wisdom

[Press et al.] "Numerical Recipes"

[Atkinson 85]

[Ralston & Rabinowitz 78]

Classical Stuff

[Acton 70]
- Numerical integration and interpolation are perhaps the two earliest goals of numerical computation.

- Generally speaking, integration is a smoothing operation, & is relatively insensitive to noise, and forgiving.

- Interpolation is less forgiving
- Differentiation is least forgiving, ... etc.

Our path for the next three assignments is through three central applications of integration:

*quadrature* $\rightarrow$ ode's $\rightarrow$ pde's

**Strategy for Numerical Integration:**

\[ \int_a^b f(x)\,dx = \text{area under curve} \ f(x) \]

1. Break up interval into pieces of width $h$

2. Approximate $f(x)$ by a polynomial over a few intervals (degree $n$, $n+1$ points, $n$ intervals)

3. Integrate these polynomial approximations & add up.

[ 4. Repeat for smaller $h$ until convergence ]
Standard Simple Method: degree 1 polynomials

Trapezoidal Rule

2 points, 1 interval

\[ \int_a^b f(x) \, dx \approx \text{area under trapezoid} \]
\[ \lambda = \frac{h}{2} (f_a + f_b) \]

"Extended" Version:

\[ \begin{align*}
\int_a^b f(x) \, dx &= \frac{h}{2} \left( \frac{1}{2} f_a + \frac{1}{2} f_{a+1} \right) + \frac{h}{2} \left( \frac{1}{2} f_{a+1} + \frac{1}{2} f_{a+2} \right) + \cdots \\
&= h \cdot \left( \frac{1}{2} \frac{1}{2} \frac{1}{2} \cdots \frac{1}{2} \right) \left[ \text{weights} \right]
\end{align*} \]

→ Don't recompute old ordinates when \( h \) is halved

\[ h \text{ sum}_i = \text{solid values, endpoints weighted by } \frac{1}{2} \]
\[ \frac{h}{2} \text{ sum}_{i+1} = \text{sum}_i + \text{dotted values} \]

\[ \text{save } \frac{h}{2} \]
What's wrong with

\[
\pi_{i+1} = \pi_i + h
\]

Do

\[
\pi_k = (\text{Interval length}) \times k/N \quad k = 0, \ldots, N
\]

Simpson's Rule

the next higher order polynomial

quadratic, 3 points, 2 intervals

\[
\int_{a}^{a+2h} f(x) \, dx \approx \text{area under quadratic}
\]

\[
= \frac{h}{3} \left[ f(a) + 4f(a+h) + f(2a+2h) \right]
\]

Composite ("extended") Version:

\[
\frac{h}{3} \begin{bmatrix}
1 & 4 & 1 \\
4 & 1 & 4 \\
1 & 4 & 1 \\
\end{bmatrix}
\]

\[
\frac{h}{3} \begin{bmatrix}
1 & 4 & 2 & 4 & 2 & 4 & \cdots & 4 & 1 \\
\end{bmatrix}
\]

Weights
Re-using points in Simpson's Rule is slightly trickier:

\[ \frac{3}{h} A_n = 1 \quad 4 \quad 2 \quad 4 \quad 2 \quad 4 \quad \ldots \quad 4 \quad 1 \]

\[ = \frac{1}{B_n} \left( \sum \frac{C_n}{2} \right) + 4 \cdot \text{sum of new ordinates} \]

\[ \frac{3}{(h/2)} A_{n+1} = \left[ B_n + \frac{C_n}{2} \right] + 4 \cdot \text{sum of new ordinates} \]

Start with

\[ \frac{3}{(b-a)/2} A_1 = \frac{1}{B_1} \quad 4 \quad C_1 \quad a \quad b \]

Note that Numerical Recipes [Press et al.] shows how to implement Simpson's Rule using Trapezoidal Rule. (E.g. 4.2.4, etc.)

Additional Refinement: Romberg quadratures: [Acton 96]

<table>
<thead>
<tr>
<th>( k )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h/2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h/4 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h/8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|   | \( T_{00} \) | \( R_{11} \) | \( R_{12} \) | \( R_{13} \) | \( R_{22} \) | \( R_{23} \) | \( R_{33} \) |

Computed from \( T_5 \)’s

\[ d = \frac{4 \cdot c - 5}{4 \cdot b - 1} \]

Very accurate & efficient if integral reasonably well behaved.
Classical Error Analysis

In one interval:

\[ \int_{a}^{b} f(t) \, dt = \frac{h}{2} (f_a + f_b) + E \quad \text{sum} \]

Use Taylor's Series with remainder:

\[ f_t = f_a + (t-a) f'_a + \frac{(t-a)^2}{2!} f''_a \]

for some \( a \leq \theta \leq b \)

\[ \text{LHS} = \int_{a}^{b} \left[ 1 \, dt = (b-a) f_a + \frac{(b-a)^2}{2} f'_a + \frac{(b-a)^3}{3!} f''_a \right] \]

\[ = \text{RHS} = \frac{h}{2} \left[ f_a + f_a + (b-a) f'_a + \frac{(b-a)^2}{2} f''_a \right] + E \]

\[ \frac{h}{2} f_a + \frac{h^2}{6} f'_a + \frac{h^3}{4} f''_a = \frac{h^3}{6} f_a + \frac{h^2}{2} f'_a + \frac{h^3}{4} f''_a + E \]

\[ \Rightarrow \quad E = - \frac{h^3}{12} f''_a \quad \theta \in [a,b] \]

Extended Rule

\[ E_{\text{TOTAL}} = - \frac{h^2}{12} \left[ \sum_{i} h f_{\theta_i}'' \right] = - \frac{h^2}{12} (b-a) f''_a \quad \theta \in [a,b] \]

(Uses mean value theorem, see [Atkinson 85])
this predicts that as we continue to halve $h$, 

$$E_{\text{new}} \approx \frac{1}{4} E_{\text{old}}$$

* The derivation assumes $f$ is sufficiently smooth, that higher derivatives used are continuous.

Estimate [Atkinson 85]:

$$E_{\text{TOTAL}} = -\frac{h^2}{12} \left[ \sum_{i=1}^{\text{intervals}} h f' \theta_i \right]$$

$$\xrightarrow{h \to 0} \int_a^b f''(x) dx = f'(b) - f'(a)$$

$$E_{\text{TOTAL}} \approx -\frac{h^2}{12} \left[ f'(b) - f'(a) \right]$$

Two implications:

1) can be much better if $f'(b) = f'(a)$

2) Suggests correcting result by this estimate.
Corresponding Result for Simpson's Rule

\[ E_{\text{TOTAL}} = -\frac{h^4}{90} (b-a) f^{(iv)}(\theta) \quad \theta \in [a, b] \]

\[ \approx -\frac{h^4}{90} \left[ f''(b) - f''(a) \right] \]

For smooth enough functions, the error is decreased by about \( \frac{1}{2^4} = \frac{1}{16} \) each halving of \( h \).

**More Advanced Techniques:**

- **Gauss quadrature** uses non-equispaced points.
  - Much more complicated!
  - Much more accurate in appropriate cases.

- **"Open formulas"**
  - Uses only internal points.
Common Problems that require some thought:

\[ \int_{A}^{\infty} p(x) \, dx \quad \text{infinite limit} \]

\[ \implies \text{I want you to deal with this in Assignment 2. Think!} \]

"Sick" integrals (strongly recommend [Acton 96])

E.g.,

\[ I = \int_{0}^{\pi/2} \frac{\cos x}{\sqrt{x}} \, dx \]

But nothing really bad is happening near \( x = 0 \), because

\[ \int_{0}^{\pi/2} \frac{1}{\sqrt{x}} \, dx \text{ is finite.} \]

The infinity can be removed; let \( x = u^2 \), \( dx = 2u \, du \)

\[ I = \int_{0}^{\pi/2} \frac{\cos(u^2)}{u} \cdot 2u \, du = 2 \int_{0}^{\pi/2} \cos(u^2) \, du \]

\( \sqrt{\text{& logs}} \) can often be fixed this way if they produce an infinity.
Examples of error behavior (Atkinson 85)

\[ I^{(1)} = \int_0^1 e^{-x^2} \, dx \approx 0.74682413281234 \]

<table>
<thead>
<tr>
<th># pts.</th>
<th>Error - Trapezoidal</th>
<th>Ratio</th>
<th>Error - Simpson's</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.55E-2</td>
<td></td>
<td>-8.56E-4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.84E-3</td>
<td>4.02</td>
<td>-3.12E-5</td>
<td>11.4</td>
</tr>
<tr>
<td>8</td>
<td>9.59E-4</td>
<td>4.01</td>
<td>-1.99E-6</td>
<td>15.7</td>
</tr>
<tr>
<td>16</td>
<td>2.40E-4</td>
<td>4.00</td>
<td>-1.25E-7</td>
<td>15.9</td>
</tr>
<tr>
<td>32</td>
<td>5.99E-5</td>
<td>4.00</td>
<td>-7.79E-9</td>
<td>16.0</td>
</tr>
</tbody>
</table>

\[ I^{(3)} = \int_0^{2\pi} \frac{dx}{2 + \cos(x)} = \frac{2\pi}{\sqrt{3}} \approx 3.6275987284684 \]

<table>
<thead>
<tr>
<th># pts.</th>
<th>Error - Trapezoidal</th>
<th>Ratio</th>
<th>Error - Simpson's</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-5.61E-1</td>
<td></td>
<td>-1.26</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-3.76E-2</td>
<td>14.9</td>
<td>1.77E-1</td>
<td>-9.2</td>
</tr>
<tr>
<td>8</td>
<td>-1.93E-4</td>
<td>195</td>
<td>1.23E-2</td>
<td>11.2</td>
</tr>
<tr>
<td>16</td>
<td>-5.17E-9</td>
<td>37600</td>
<td>6.43E-5</td>
<td>191</td>
</tr>
<tr>
<td>32</td>
<td>* (machine limit)</td>
<td></td>
<td>1.71E-9</td>
<td>37600</td>
</tr>
<tr>
<td>64</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

\[ I = \int_0^1 \sqrt{x} \, dx = 2/3 \]

<table>
<thead>
<tr>
<th># pts.</th>
<th>Error - Simpson's</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.860E-2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.012E-2</td>
<td>2.82</td>
</tr>
<tr>
<td>8</td>
<td>3.587E-3</td>
<td>2.83</td>
</tr>
<tr>
<td>16</td>
<td>1.268E-3</td>
<td>2.83</td>
</tr>
</tbody>
</table>
| 32      | 4.485E-4          | 2.83  | ~ bad accuracy...

\[ \text{why? so bad} \]