سؤال ۱)

خطای مدل: در این رابطه، تمام جرم سیب در یک نقطه متمرکز در نظر گرفته می‌شود و از ابعاد و شکل ظاهری آن صرف‌نظر می‌شود در حالی که ابعاد سیب می‌تواند روی مقاومت هوا تأثیر بگذارد.

همچنین از چرخش سیب و ممان اینرسی (گشتاور لختی) آن صرف‌نظر شده است.

نهایتاً از اثرات نسبیتی چشم‌پوشی شده که در مسئلهٔ سقوط سیب با توجه به سرعت و ابعاد جسم، ناچیز است.

خطای اندازه‌گیری: ابزارهای اندازه‌گیری ترازو، سرعت‌سنج یا نیروسنج همگی دقت ثابتی دارند و نمی‌توانند مقدار را با بی‌نهایت رقم اعشار گزارش کنند. پس در اندازه‌گیری پارامترها و گزارش آن خطا داریم. به عنوان مثال در یک ترازوی دیجیتال با دقت گرمی، ± 1 گرم خطا داریم.

خطای گرد کردن: اگر بخواهیم نتیجهٔ ضرب کردن $m\times a$ را با تعداد ارقام مشخصی گزارش کنیم باید حاصل را گرد کنیم. اگر به روش symmetric تا t رقم گرد کنیم، خطای مطلق حدی آن $\frac{1}{2}\times 10^{-t}$ است.

خطای عملیات: در عملیات ضرب $m\times a$ انتشار خطا داریم. خطای حاصل‌ضرب از رابطهٔ زیر به‌دست می‌آید:

$$e_{\bar{m}\times\bar{a}} \leq \bar{m}\, e_{\bar{a}} + \bar{a}\, e_{\bar{m}}$$

سوال دوم

الف)

$$\Delta = b^2 - 4ac = 11.01^2 - 4 \times 1.002 \times 0.01265 = 121.22 - 0.05 = 121.17$$

$$\rightarrow \sqrt{\Delta} = \sqrt{121.17} \simeq 11.01$$

$$\rightarrow \begin{cases} \bar{x_1} = \dfrac{-b+\sqrt{\Delta}}{2a} = \dfrac{-11.01+11.01}{2.004} = 0 \\ \bar{x_2} = \dfrac{-b-\sqrt{\Delta}}{2a} = \dfrac{-11.01-11.01}{2.00} = -\ 11.01 \end{cases}$$

این در حالی است که ریشه‌های اصلی مطابق زیر می‌باشند:

$$\begin{cases} x_1 = \dfrac{-b+\sqrt{\Delta}}{2a} = \dfrac{-11.01+11.00076}{2.004} = -\ 0.0011 \\ x_2 = \dfrac{-b-\sqrt{\Delta}}{2a} = \dfrac{-11.01-11.00076}{2.004} = -\ 10.9869 \end{cases}$$

در نتیجه خطای نسبی هر یک از ریشه‌ها مطابق زیر خواهد بود:

$$\begin{cases} \delta_{x_1} = \dfrac{|x_1 - \bar{x_1}|}{x_1} = \dfrac{|0-0.0011|}{0.0011} = 1 \\ \delta_{x_2} = \dfrac{|x_2 - \bar{x_2}|}{x_2} = \dfrac{|11.01-10.9869|}{10.9869} = 0.002 \end{cases}$$

ب)

$$\delta_{x_1} = \frac{|x_1 - \bar{x_1}|}{x_1}$$

با توجه به آن که مطابق فرمول بالا، در خطای نسبی ریشه اول، مقدار عددی دو متغیر $\bar{x_1}$ و x_1 بسیار به هم نزدیکند، خطای نهایی تفریق این دو مقدار بسیار زیاد شده‌است. به عبارتی دیگر، در محاسبه خطای نسبی تفریق دو متغیر، اگر مقدار آن دو تقریبا برابر یک‌دیگر باشند، خطای نسبی افزایش خواهد یافت.

ج)

مطابق فرمول زیر، در محاسبه ریشه اول، صورت کسر بسیار نزدیک به صفر بود.

$$\bar{x_1} = \frac{-b+\sqrt{\Delta}}{2a} = \frac{-11.01+11.01}{2.004} = 0$$

برای حل این مشکل کافی است مکمل صورت را در صورت و مخرج کسر ضرب کنیم:

$$\bar{x_1} = \frac{(-b+\sqrt{\Delta}) \times (-b-\sqrt{\Delta})}{2a \times (-b-\sqrt{\Delta})} = \frac{4ac}{2a \times (-b-\sqrt{\Delta})} = \frac{0.05}{2.004 \times (-11.01-11.01)} = -\ 0.0011$$

با توجه به آن که این مشکل برای ریشه دوم نبود، نیازی به اجرای این روند برای آن ریشه نمی‌باشد.

سؤال ۳)

از دو اتحاد مثلثاتی زیر استفاده می‌کنیم:

$$\sin(2\alpha)=2\sin(\alpha)\cos(\alpha)$$
$$1-\cos(2\alpha)=2\sin^2(\alpha)$$

تابع f به صورت زیر بازنویسی می‌شود:

$$f(x)=\frac{1-\sin(x)}{\sin(x)}=\frac{2\sin^2(x/2)}{2\sin(x/2)\cos(x/2)}=\frac{\sin(x/2)}{\cos(x/2)}=\tan(x/2)$$

بنابر سری تیلور تابع تانژانت داریم:

$$f(x)=\tan(x/2)\simeq(x/2)+\frac{(x/2)^3}{3}+\frac{2(x/2)^5}{15}+\frac{17(x/2)^7}{315}+\ldots=\frac{x}{2}+\frac{x^3}{24}+\frac{x^5}{240}+\frac{17x^7}{40320}+\ldots$$

در نتیجه خطای مطلق تخمین برابر است با

$$e_{g(x)}=|g(x)-f(x)|\simeq\left|\frac{x^5}{240}+\frac{17x^7}{40320}+\ldots\right|$$

(برای xهایی که بسیار به صفر نزدیک‌اند، می‌توان از جملات با درجات بالاتر صرف‌نظر کرد.)

سؤال ۲)

الف) برای خطای مطلق a_1 تا a_n داریم:

$$\bar a_1 - e_{\bar a_1} \le a_1 \le \bar a_1 + e_{\bar a_1}$$
$$\bar a_2 - e_{\bar a_2} \le a_2 \le \bar a_2 + e_{\bar a_2}$$
$$\vdots$$
$$\bar a_n - e_{\bar a_n} \le a_n \le \bar a_n + e_{\bar a_n}$$

$$(\bar a_1+\bar a_2+\dots+\bar a_n)-(e_{\bar a_1}+e_{\bar a_2}+\dots+e_{\bar a_n}) \le a_1+a_2+\dots+a_n \le (a_1+a_2+\dots a_n)+(e_{\bar a_1}+\dots+e_{\bar a_n})$$

$$\Rightarrow |a_1+a_2+\dots+a_n-(\bar a_1+\bar a_2+\dots+\bar a_n)| \le e_{\bar a_1}+e_{\bar a_2}+\dots+e_{\bar a_n}$$

$$\delta_{\bar a_1+\bar a_2+\dots+\bar a_n} \le \left|\frac{e_{\bar a_1}+\dots+\bar a_n}{\bar a_1+\dots+\bar a_n}\right| \le \left|\frac{e_{\bar a_1}+\dots+e_{\bar a_n}}{\bar a_1+\dots+\bar a_n}\right|$$

$$\le \frac{\bar a_1}{\bar a_1+\dots+\bar a_n}\times\frac{e_{\bar a_1}}{\bar a_1}+\dots+\frac{\bar a_n}{\bar a_1+\dots+\bar a_n}\times\frac{e_{\bar a_n}}{\bar a_n}=\frac{\bar a_1}{\bar a_1+\dots+\bar a_n}\delta_{\bar a_1}+\dots+\frac{\bar a_n}{\bar a_1+\dots+\bar a_n}\delta_{\bar a_n}$$

حال فرض کنیم $\delta_m=\max(\delta_{\bar a_1},\delta_{\bar a_2},\dots,\delta_{\bar a_n})$ در این صورت داریم:

$$\delta_{\bar a_1+\bar a_2+\dots+\bar a_n} \le \frac{\bar a_1}{\bar a_1+\dots+\bar a_n}\delta_m+\dots+\frac{\bar a_n}{\bar a_1+\dots+\bar a_n}\delta_m=\left(\frac{\bar a_1+\dots+\bar a_n}{\bar a_1+\dots+\bar a_n}\right)\delta_m=\delta_m$$

پس $\delta_{\bar a_1+\bar a_2+\dots+\bar a_n} \le \delta_m$

ب)

$$x = 12.3456789$$

$$\delta_{\bar x} \le 10^{-4} \quad , \quad \delta_{\bar x} = \left|\frac{e_{\bar x}}{x}\right| = \left|\frac{x-\bar x}{x}\right|$$

$$\left|1-\frac{\bar x}{x}\right| \le 10^{-4} \rightarrow -10^{-4} \le 1-\frac{\bar x}{x} \le 10^{-4} \rightarrow -10^{-4} \le \frac{\bar x}{x}-1 \le 10^{-4}$$

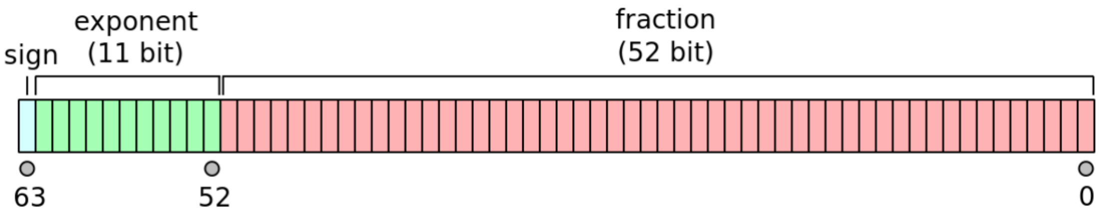
$$\rightarrow 1-0.0001 \le \frac{\bar x}{x} \le 1+0.0001 \rightarrow 0.9999\,x \le \bar x \le 1.0001\,x$$

$$\rightarrow 12.34444423321 \le \bar x \le 12.3469134679$$

طبق محاسبات بالا می‌بینیم که ارقام ۱۲٫۳۴ قطعاً درست محاسبه شده‌اند. اما درباره‌ی سایر ارقام نمی‌توان نظر داد

سوال ۵)

فرمت عدد IEEE double precision به صورت زیر است :



که در این صورت، عدد نشان‌داده شده به صورت زیر محاسبه می‌شود :

$$(-1)^{sign}(1.b_{51}b_{50}\ldots b_0)_2 \times 2^{e-1023}$$

بنابراین در عدد داده شده در صورت سوال، بخش‌های مختلف به صورت زیر است :

Number =>11000000101001010011000...00

Sign => 1

Exponent => 10000001010 = $(1034)_{10}$

Mantissa => 010100110000...00 = $(0.32421875)_{10}$

که می‌توانیم این عدد را به صورت $-1 \times (1.10000001010)_2 \times 2^{1034-1023}$ یا $-1.32421875 \times 2^{11}$ نشان دهیم. (عدد اصلی در مبنای ۱۰ برابر با 2712- می‌شود.)

حال می‌خواهیم نزدیک‌ترین عدد دهدهی بزرگتر و کوچکتر از این عدد را حساب کنیم. برای این کار کافی است مقدار $10^{-52} \times 2^{11}$ از عدد کم و یا به آن اضافه کنیم. به عبارت دیگر از بخش mantissa عدد، یک واحد کم یا به آن یک واحد اضافه کنیم. بنابراین داریم،

نزدیک‌ترین عدد کوچکتر :

$$-2712 - 10^{-52} \times 2^{11} = (-2712.00000000000045474735088)_{10}$$

نزدیک‌ترین عدد بزرگتر :

$$-2712 + 10^{-52} \times 2^{11} = (-2711.99999999999954525264911)_{10}$$

سوال6)

```python
import math


def stirling(x):
    return math.sqrt(2 * math.pi * x) * math.pow((x / math.e), x)


n = float(input())
n_fact = math.gamma(n + 1)
n_almost_fact = stirling(n)
abs_err = abs(n_fact - n_almost_fact)
rel_err = abs_err / n_fact
print('absolute error:', round(abs_err, 4))
print('relative error:', round(rel_err, 4))
```