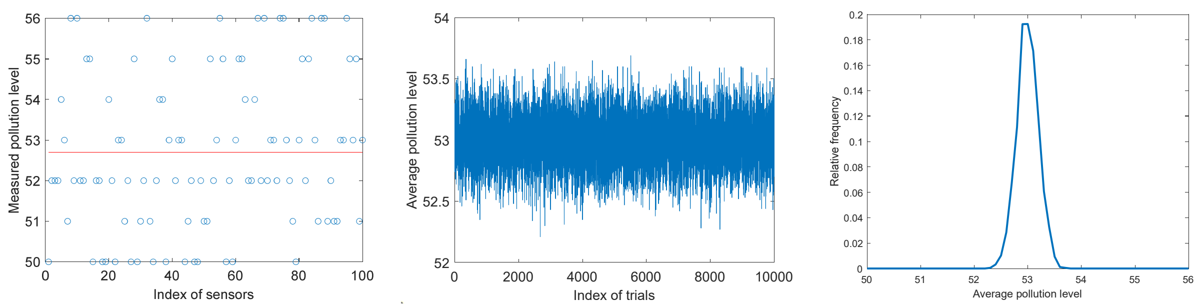
## ECE3522 Practicum 3 : Achieving More Accurate Measurements

**Penalty will be imposed after 4/19/20 10:00 pm (1 point each hour; fraction rounds up)**

A certain type of sensor nodes are low-cost devices that measure the environment pollution level with a low accuracy. To improve the accuracy, $n$ sensor nodes are distributed in an area to measure the pollution level, and the results show that the outcome of each sensor node, $X_i,\ i = 1, \ldots, n$, are integers that follow independent and identically distributed (i.i.d.) discrete uniform distribution between 50 and 56. The average of the measured pollution level of $n$ such sensor nodes is expressed as $M_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$.

The attached Matlab example computes the average pollution level of 100 sensor nodes, and 10,000 independent trials are repeated to generate a histogram of the average pollution level. The results are depicted in the following plots.




(a) (b) (c)

(a) Measured pollution level at 100 sensors (the red line shows the average value)
(b) Average pollution level for 10,000 independent trials
(c) Histogram of the average pollution level

This practicum requires the following results to be computed and reported:

**Part 1.** For each of the $n$ values between 1 and 100, analytically compute the expected value and the standard derivation of $M_n$. The results must be properly derived or justified. Plot the expected value and the standard deviation with respect to $n$, and report the specific values in the following table. Describe in what relationship the standard derivation varies with $n$. [Note the variance of random variable $X$ following discrete uniform $(a\ ,\ b\ )$ is given as $\mathrm{Var}[X] = (b-a)(b-a+2)/12$.]

| $N$ | 1 | 10 | 100 |
|---|---|---|---|
| Expected value of $M_n$ ($\mathrm{E}[M_n]$) | | | |
| Standard deviation of $M_n$ ($\sigma_{M_n}$) | | | |

**Part 2.** Based on the Central Limit Theorem, when $n$ is sufficiently large, $M_n$ follows the Gaussian (normal) distribution with the expected value and standard deviation obtained in Part 1. Given $n$ = 100 and 1,000, analytically compute the probability $P[M_n \geq 53.2]$.

**Part 3.** If we want to achieve a standard deviation of 0.1 or less from the average of the measured pollution level of $n$ sensor nodes, how many sensor nodes are needed?

**Part 4.** Experimentally verify the results in Part 2 by computing the ratio between the number of trials with pollution level higher than 53.2 and the total number of trials. Comments if the results are close to those reported in Part 2. Comment if they differ.

**Part 5.** Experimentally verify the results in Part 3. In this case, the analytical probability that the average pollution level is higher than $53 + 2\sigma_{M_n} = 53.2$ is $P[M_n \geq 53.2] = 1 - \Phi(2)$ =0.02275. Using the value of $n$ obtained in Part 3, compute the ratio between the number of trials with pollution level higher than 53.2 and the total number of trials. Comment if this value differs to $P[M_n \geq 53.2]$ =0.02275.

**Submit by uploading to Canvas or e-mail to ece3522.temple@gmail.com**

(a) A single Word or pdf file containing materials in the following order:

- Your name, TUid
- Clear label each of the five parts and report the required values, figures, and observations. All plots must be properly labelled.
- Key references (website, book, paper, or name of collaborated students) if any
- Matlab codes

(b) All Matlab codes in .m format

**Important: The Matlab code must be included in the Word or pdf file report, AND as a separate attachment.** The name of the Word of pdf file must be in the following format with a proper file extension: Lastname_Firstname _practicum3_ECE3522 .

**Requirements**

Each student should complete the practicum separately. No group work is allowed.

A student can seek help, but the report must be his/her work and thus he/she must understand every word and every line of Matlab code reported. Copy and paste the Matlab codes from others work or from online resources are not permitted.

**Appendix**

```
N_Sensors = 100;                % number of sensors in each trial
N_Trials = 10000;               % number of trials
range = [50 56];                % range of measured pollution level
hist_range = [50:.1:56];        % range of histogram plot

for iloop = 1:N_Trials
    score = randi(range, N_Sensors, 1);
    score_ave(iloop) = sum(score)/N_Sensors;
end

hist_result = hist(score_ave, hist_range);

figure(1)
plot(score,'o')
hold on
plot(score_ave(end)*ones(N_Sensors,1),'r');
xlabel('Index of sensors'); ylabel('Measured pollution level')

figure(2)
plot(score_ave)
xlabel('Index of trials'); ylabel('Average pollution level')

figure(3)
plot(hist_range, hist_result/N_Trials,'-','LineWidth',2)
xlim(range)
xlabel('Average pollution level'); ylabel('Relative frequency')
```