ME 103 Experimentation and Measurements
Spring 2026 George Anwar

HW2

Due: Monday, February 23rd at 11:59 pm

- This homework will cover the hypothesis testing and an introduction to frequency response methods and spectral analysis.
- In all of the questions, **show your work**, not just the final answer. Unless we explicitly state otherwise, you may expect full credit only if you explain your work succinctly, but clearly and convincingly. For coding questions, attach a screenshot of your code and output.
- Present your answers with **a suitable number of significant figures** for each question. Show your work, including a mathematical formula or the MATLAB or Python code you wrote, before reaching the result. You may need to install the Statistics Toolbox if using MATLAB.
- Throughout this assignment, neglect systematic (bias) errors. Also, assume a normal distribution for the underlying distribution (population) if necessary.
- If you have a confirmed disability that precludes you from complying fully with these instructions or with any other parameter associated with this problem set, please alert us immediately about reasonable accommodations afforded to you by the DSP Office on campus.
- **Start early. Some of the material is prerequisite material not covered in lecture; you are responsible for finding resources to understand it.**

Deliverables

Submit a PDF of your homework to the **Gradescope assignment** entitled "{Your Name} HW2". You **must** typeset your homework in LaTeX (submit PDF format, not .doc/.docx format) . Mac Preview, PDF Expert, and FoxIt PDF Reader, among others, have tools to let you sign a PDF file. We want to make *extra clear* the consequences of cheating.

0 Honor Code

I will adhere to the Berkeley Honor Code: specifically, as a member of the UC Berkeley community, I act with honesty, integrity, and respect for others. Failure to comply with these guidelines can be considered an academic integrity violation. Please email Professor Anwar ganwar@berkeley.edu if you have any questions!

- **List all collaborators. If you worked alone, then you must explicitly state so. Read the following statement and sign below if you agree**:

  "I certify that all solutions in this document are entirely my own and that I have not looked at anyone else's solution. I have given credit to all external sources I consulted."

  Signature : ______________________ *Date* : ______________________

  While discussions are encouraged, *everything* in your solution must be your (and only your) creation. Furthermore, all external material (i.e., *anything* outside lectures and assigned readings, including figures and pictures) should be cited properly. We wish to remind you that consequences of academic misconduct are *particularly severe*!

- Violation of the Code of Conduct will result in a **zero** on this assignment and may also result in disciplinary action.

1 Hypothesis Testing Basics [18 points]

(a) [9 pts] A food manufacturer claims its ready-meals contain on average no more than 2.0g of salt per serving. From a sample of 30 servings the mean salt content is found to be 2.2g. You may assume that the variance is known to be 0.2. Use an appropriate hypothesis test to determine whether the claim is correct. You may use a significance level of 2.5%.

Solution: Let μ be the true mean salt (g) per serving of the ready-meals. Then define our hypothesis as follows

$$H_0 : \mu \leq 2.0$$
$$H_a : \mu > 2.0$$

We then have the following assumptions. We have a random sample of $n = 30$ servings with a known population variance of 0.2. Therefore, n is large enough for the central limit theorem to hold. Therefore, we perform a **one-sample z-test** for a population mean with the following test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.2 - 2.0}{\sqrt{0.2/30}} = 2.449$$

We now obtain the p-value. Since H_a points to greater than, large positive z-values are the evidence against H_0 implying a right-tailed test. This is a one-tailed test so the critical value at $\alpha = 0.05$ is 1.96.

$$p = 1 - \Phi(2.449) = 1 - 0.9929 = 0.0071$$

Therefore, at $\alpha = 0.025$, $p = 0.0071 < 0.025$. Also, $z = 2.449 > z_{0.975} = 1.96 \implies$ **reject H_0**. Thus, there's statistically significant evidence that the true mean salt content exceed 2.0g per serving; the manufacture's "no more than 2.0g" claim is not supported at the 2.5% level.

(b) [9 pts] 26 students who attended a course in-person obtained an average score of 60% in an exam, the standard deviation of their scores was 10%. There were an additional 36 students who attended online only, these students obtained an average score of 55% with a standard deviation of 12%. Is there a difference in the average exam scores between these two groups? You may use a significance level of 5.0% and you may assume equal variance.

Solution: Let $\mu_{in-person} - \mu_{online}$ represent the true difference in mean exam scores. Then define our hypothesis as follows

$$H_0 : \mu_{in-person} - \mu_{online} = 0$$
$$H_a : \mu_{in-person} - \mu_{online} \neq 0$$

We then have the following assumptions. We have two independent random samples, with different population variances. We do not meet the requirements to apply the central limit theorem. Therefore, we apply a **Welch's two-sample t-test** for a difference of means with **pooled variance**.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(26 - 1) \times 100 + (36 - 1) \times 144}{60} = 125.\bar{6}$$

The t-statistic is then given by

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{60 - 55}{\sqrt{125.\bar{6}}\sqrt{\frac{1}{26} + \frac{1}{36}}} = 1.733$$

We have a degree of freedom of $n_1 + n_2 - 2 = 60$. At $\alpha = 0.05$, the critical $|t| > t_{0.975,60} = 2.00$. Since $1.733 < 2.00$, fail to reject H_0. Therefore there isn't sufficient evidence at the 5% level to conclude that the mean exam scores differ between in-person and online students.

2 Non-Parametric Statistics [19 points]

(a) [3 pts] Briefly describe two advantages and one disadvantage of using non-parametric tests. **Solution:** Multiple advantages/disadvantages possible, for example: **Advantages**: apply to ordinal data (i.e. type of degree you have), normality not required. **Disadvantages**: if we can assume normality then tests that leverage this knowledge are generally more powerful.

(b) [8 pts] A business contracts two parcel delivery companies (Company A and Company B) to deliver its goods. This business obtains feedback about how satisfied each customer was with their delivery. Satisfaction is rated on a scale of 0 to 20, the higher the rating the more satisfied the customer.

Company	**Satisfaction**
B	9
A	10
A	11
B	15
A	17
B	19
B	20
A	20

Is there a difference in the median customer satisfaction between the deliveries made by these two companies? Use an appropriate non-parametric test to answer this question. You may use a significance level of 5%.

Solution: Let the difference in population locations (medians) of satisfaction for Company A vs B be the parameter of interest. We can then define our hypothesis as follows

$$H_0 : \text{The probability distributions are identical}$$
$$H_a : \text{Probability distribution for A is shifted either to the left/right of that for B}$$

Our assumptions are that we have two independent random samples; ordinal/continuous scale; similar shape/dispersion across groups so a location shift is meaningful; ties handled by average ranks. We can then apply the **two-sided Wilcoxon rank-sum test** or the **Mann-Whitney U**.

The data has been sorted by satisfaction already, hence the ascending ranks corresponding to the sorted scores will be $\{1, 2, 3, 4, 5, 6, 7, 8\}$. Note that the two 20's are tied for ranks 7 and 8 so we assign them the average $(7 + 8/2 = 7.5)$. Therefore we can calculate the sum of the ranks of company A and B as

$$T_A = 2 + 3 + 5 + 7.5 = 17.5, \quad T_B = 1 + 4 + 6 + 7.5 = 18.5$$

Since $n_A = n_B = 4$, the test statistic T can be either T_A or T_B and $T_L = 11$, $T_U = 25$. The rejection region is if $T_1 \geq T_U$ (or $T_1 \leq T_L$); $T_2 \leq T_U$ (or $T_2 \geq T_U$). Since 17.5 or 18.5 is not ≤ 11 or ≥ 25, we have insufficient evidence to reject the null hypothesis that the distributions are identical.

(c) [8 pts] Six students rate two statistics textbooks on a scale from 1 to 10, where 1 is the easiest to understand and 10 being the hardest. The books are *All of Statistics* and *Bayesian Data Analysis*. Use an appropriate non-parametric test to determine whether students find the book *All of Statistics* harder to understand than *Bayesian Data Analysis*. You may use a significance level of 5%.

Student	**All of Statistics**	**Bayesian Data Analysis**
1	3	9
2	5	1
3	2	4
4	7	2
5	1	8
6	8	3

Solution: Let us first denote "All of Statistics" as Book A, and "Bayesian Data Analysis" as Book B. Let the parameter of interest be the median paired difference in hardness ratings. We can then define our hypothesis as follows

$$H_0 : \text{The ratings are identical}$$
$$H_a : \text{Book A is rated higher (i.e. harder) than Book B}$$

We make the assumptions that the data from the same 6 students is paired; differences are symmetrically distributed about their median; measurement is at least ordinal; no zero differences (none here). We then use the **one-sided Wilcoxon Matched-Pairs Signed Rank Test**.

Let's start by computing the differences $d_i = A_i - B_i$ and then rank the absolute differences (average ranks for ties)

Student	A	B	$d_i = A_i - B_i$	$\lvert d_i \rvert$	Rank
1	3	9	-6	6	5
2	5	1	$+4$	4	2
3	2	4	-2	2	1
4	7	2	$+5$	5	3.5
5	1	8	-7	7	6
6	8	3	$+5$	5	3.5
			$T_+ = \sum$ (positive ranks)		9
			$T_- = \sum$ (negative ranks)		12

Then by summing the positive and negative ranks we get

$$T_+ = 2 + 3.5 + 3.5 = 9, \quad T_- = 5 + 1 + 6 = 12$$

For a one-tailed test with $n = 6$, $\alpha = 0.05$, the critical value is 2. With the H_a being that the median is greater than 0, we should check the large T_+ or equivalently a small T_-. Here T_+ is not large enough and T_- is not small enough, so we fail to reject the null hypothesis.

3 Hypothesis Testing for two samples [22 points]

Two different formulations of an elastomeric material are being evaluated, and Young's moduli at small strains are measured from a sample of pieces of each material. One formulation ('Treatment 1') results in a sample mean of 2.085 MPa while the other ('Treatment 2') results in a sample mean of 2.071 MPa. The sample standard deviations are 25 kPa and 34 kPa respectively.

As you carry out the following computations, show your working for each part and summarize your final computed results in the table on the next page (a Word version is available on bCourses if you wish to edit it).

(a) [2 points] Consider first the case (Case A in the table) where the above data are obtained from sample sizes of 15 measurements for each treatment. Calculate the effective number of degrees of freedom using the following formula:

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$$

Solution: Plugging into the above formula with the given values and rounding down to the nearest integer gives 25 degrees of freedom.

(b) [2 points] Calculate the t-statistic for the difference between the sample means, i.e.:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Solution: Plugging into the above formula gives a t-statistic of 1.284818.... = 1.285 (4 s.f.) or 1.3 (2 s.f.) The sample standard deviation is given to 2 s.f., so 2 s.f. for the t statistic is probably the most appropriate, although it would also be OK not to round at this stage and keep the full precision of the calculation up to the point where the t-statistic is compared with the critical t-value.

(c) [2 points] Using tables or software, find the critical t-values for:

(i) A 2-tailed t-test at the 5% significance level where the null hypothesis is $H_0 : \mu_1 = \mu_2$ and the alternate hypothesis is $H_a : \mu_1 \neq \mu_2$. The values μ_1 and μ_2 are the population mean Young's moduli resulting from treatments 1 and 2 respectively.

Solution: Using (in Excel) `tinv(0.05,25)` where 0.05 is the significance level of the 2-tailed test and 25 is the number of degrees of freedom, we obtain

$$t_{\text{crit}} = 2.060\ldots$$

(ii) A 1-tailed t-test at the 5% significance level where the null hypothesis is $H_0 : \mu_1 = \mu_2$ and the alternate hypothesis is $H_a : \mu_1 > \mu_2$.

Solution: Using (in Excel) `tinv(0.10,25)`, we obtain

$$t_{\text{crit}} = 1.706\ldots$$

Note that the Excel `tinv()` formula is for a 2-tailed test, so we double the significance level value if we want the critical value for a 1-tailed test (since the area of the tail needs to be twice as large for a 1-tailed test).

(Please note that Excel also has `t.inv()` [note the dot] and `t.inv.2t()` functions, which behave differently. Please see the documentation for the differences.)

(d) [2 points] Evaluate the two hypothesis tests described in 1.3(a) and 1.3(b): is the null hypothesis accepted or rejected in each case?

Solution: In both cases, the t-statistic from the data is less than the critical t value, so the null hypothesis is failed to be accepted in both case: there is no evidence at the 5% level of a difference between the population means for the two treatments.

(e) [6 points] Now assume that the sample means and standard deviations are still as given above, but instead they were calculated from smaller and unequal sample sizes: $n_1 = 7$ for treatment 1 and $n_2 = 5$ for treatment 2 (this is Case B in the table below). Recalculate the effective degrees of freedom and t-values and re-evaluate the two hypothesis tests (1- and 2-tailed) assuming these smaller sample sizes.

Solution: Please see the table below for the calculations. Again, both null hypotheses are accepted.

(f) [2 points] Comment on the impact of the sample sizes on the conclusions that are drawn in the hypothesis tests.

Solution: We note that when the sample sizes are smaller, but the sample means and standard deviations remain the same, the t-statistic for the data goes down, because the standard error of the mean is larger and the difference in sample means is a smaller multiple of the standard error. On the other hand, the critical t values become larger in magnitude when there are fewer degrees of freedom: this is because the tails of the t distribution get thicker as the number of degrees of freedom reduces, so a higher t value is required to reach the point where the tail above that t-value has a given area (encloses a given probability). Hence, reducing sample size tends to reduce the magnitude of the t statistic but increase the magnitude of critical t values. Although, in this example, none of the null hypotheses was rejected, and the conclusions were therefore the same in all cases, we came closer to rejecting the null hypothesis in the larger sample size case.

(g) [1 point] Fill in the summary table on the next page.

	Treatment 1	**Treatment 2**
Sample mean, $\bar{x}$ (MPa)	2.085	2.071
Sample standard deviation, S_x (MPa)	0.025	0.034
Case A		
Sample size	15	15
Effective degrees of freedom	25	
t statistic calculated for data	1.285	
t_{crit}, 2-tailed test where $H_0 : \mu_1 = \mu_2$ and $H_a : \mu_1 \neq \mu_2$	2.060	
Accept or reject null hypothesis for 2-tailed test?	Failed to reject	
t_{crit}, 1-tailed test where $H_0 : \mu_1 = \mu_2$ and $H_a : \mu_1 > \mu_2$	1.708	
Accept or reject null hypothesis for 1-tailed test?	Failed to reject	
Case B		
Sample size	7	5
Effective degrees of freedom	6	
t statistic calculated for data	0.782	
t_{crit}, 2-tailed test where $H_0 : \mu_1 = \mu_2$ and $H_a : \mu_1 \neq \mu_2$	2.447	
Accept or reject null hypothesis for 2-tailed test?	Failed to reject	
t_{crit}, 1-tailed test where $H_0 : \mu_1 = \mu_2$ and $H_a : \mu_1 > \mu_2$	1.943	
Accept or reject null hypothesis for 1-tailed test?	Failed to reject	

4 Nyquist criterion and frequency resolution [10 points]

(a) [2 points] First, suppose that it can be confidently assumed that there is no frequency content in the accelerometer's output signal that is above 2 kHz. What is the minimum sample rate that should be used for the measurements?

Solution: According to the Nyquist criterion, a sampling rate/frequency of at least 2×2 kHz $=$ 4kHz is needed.

(b) [2 points] To reliably test the model of the machine's vibration, the peak resonant frequency must be resolved to ±1 Hz. How can this level of resolution be achieved? How long will the measurement take to complete if the minimum number of samples is recorded?

Solution: The frequency resolution is the sampling frequency divided by the number of samples taken. In this case therefore we require:

$$\Delta f = 1Hz = \frac{f_s}{N} = \frac{4kHz}{N}$$

So the number of required samples is $N > 4000$. The time taken to acquire this number of samples is 1 second (number of samples multiplied by sampling period).

(c) [5 points] Next, suppose that the accelerometer may be sensing some additional harmonics above 2 kHz. To avoid the possibility of aliasing, a first-order low-pass filter is connected between the accelerometer and the A2D board. The -3 dB frequency of the filter is set to 2.5 kHz. The filter can be assumed to cut off signals completely if they occur at frequencies where the gain of the filter is 0.05 or less. What should the sampling rate now become to avoid aliasing?

Solution:

First, find the frequency (call it $f_{0.05}$) at which the magnitude (gain) of the transfer function of the first-order filter falls to 0.05. Since the $-3\,\mathrm{dB}$ frequency is $2.5\,\mathrm{kHz}$, we require:

$$0.05 = \frac{1}{\sqrt{1 + \left(\frac{f_{0.05}}{2.5\,\mathrm{kHz}}\right)^2}}$$

Which solves to

$$f_{0.05} = 49.9\,\mathrm{kHz}$$

The sampling frequency would need to be at least twice as large as $f_{0.05}$ to avoid aliasing of the highest frequency that might possibly appear in the signal without being completely cut off. Therefore the sampling frequency should be at least $99.9\,\mathrm{kHz}$ or about $100\,\mathrm{kHz}$!

(d) [1 point] Comment on your result in part 3 and whether there may be a better alternative to a first-order low-pass filter for this experiment.

Solution: This sampling rate is 25 times larger than the sampling frequency we started with. It shows how the very slow roll-off of a first order filter can result in quite inefficient use of sampling bandwidth.

5 Manifestation of aliasing [5 points]

A 500 Hz sine wave is sampled at a frequency of 4096 Hz. A total of 2048 points are taken.

(a) [1 point] Calculate the Nyquist frequency for this sampling process.

Solution: Nyquist frequency is half the sampling frequency, 2048 Hz

(b) [1 point] Calculate the frequency resolution of the sampling process.

Solution: Frequency resolution is 4096 Hz/2048 = 2 Hz

(c) [3 point] Suppose that the sampled waveform contains several harmonics of 500 Hz (i.e., signals at integer multiples of 500 Hz). Which of these can be accurately measured? What happens to the others?

Solution: The harmonics of 1000 Hz, 1500 Hz and 2000 Hz will be accurately measured and any higher harmonics will be aliased (i.e. appear as erroneously low frequencies below the Nyquist frequency.)

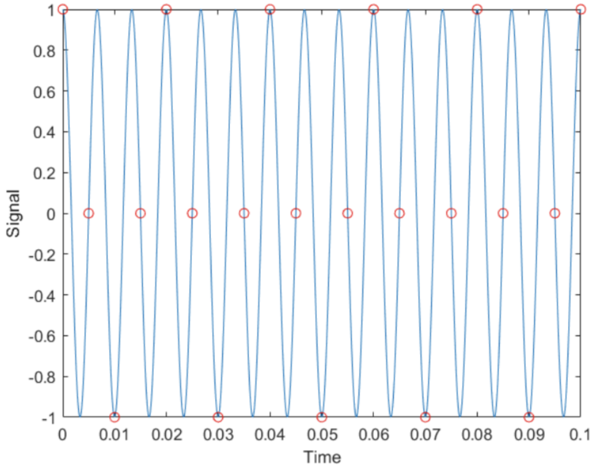
6 Manifestation of aliasing [8 points]

A 150 Hz cosine wave is sampled at a rate of 200 Hz.

(a) [4 points] Draw the wave and the temporal locations at which it is measured. You might want to use or modify the Matlab code below to accomplish this:

```
% Define duration T that is 20 sample periods
T = 20/200;
% Define time vector with 1000 points
t = linspace(0,T,1000);
% Generate cosine wave signal at frequency 150 Hz
B = cos(2*pi*150*t);
% Plot the signal waveform
plot(t,B)
% Generate time series of samples, t1
t1 = 0:(1/200):T;
% Evaluate values of 150 Hz wave at the sampling intervals
B1 = cos(2*pi*150*t1);
hold on
% Plot the samples measured
plot(t1,B1,'ro')
```

Solution:



(b) [1 point] What apparent frequency is measured?

Solution: From the graph, the period of the aliased signal (red circles) is three times longer than that of the signal being sampled; i.e. apparent frequency is 150 Hz / 3 = 50 Hz.

(c) [3 points] Provide an explanation for the numerical relationship between the actual signal frequency (150 Hz) and the measured frequency.

Hint: Look at material on heterodyning, and imagine that the 'carrier' wave has frequency 200 Hz and is multiplied by the signal wave, a sinusoid of 150 Hz. What output frequency/frequencies does the model for heterodyning predict?

Solution: The process of aliasing could be thought of as somewhat analogous to heterodyning a 150 Hz signal with a 200 Hz carrier wave. The process of sampling is often, in fact, modeled as the multiplication of the signal with a train of Dirac delta functions (delta train) at the sampling rate, and in terms of predicting frequency content, this is very similar to simply multiplying the two sinusoidal signals together (as is done in heterodyning). From the lecture slide on heterodyning, we know that when a signal wave with frequency ω_s is multiplied by a carrier wave of frequency ω_c, we end up with frequency components at $\omega_c + \omega_s$ and at $\omega_c - \omega_s$. In radio transmission, $\omega_c \gg \omega_s$, but the principle extends to where the two frequencies are of the same order of magnitude, as in this question. Hence, the $\omega_c - \omega_s$ term has frequency (200–150) Hz = 50 Hz, as seen in the graph above. The (200+150) Hz = 350 Hz term is not seen because it is out of the range of sampled frequencies.