Last Name:

First Name:

SCIPER:

Practice Final Exam – Microengineering 110

Spring 2025

1. A company is planning to purchase a robot to perform a particular manufacturing task. Three different vendors submit robots for consideration. Vendor A will sell the robot for 1 Million CHF. Vendor B will sell the robot for 1.1 Million CHF. Vendor C will sell the robot for 1.2 Million CHF.

   The company generally plans to buy the robot that completes the task most quickly. On the other hand, if there is no statistical difference in the speed of the robots, the company will buy the cheapest robot.

   The company plans to run repeated trials with each robot, and determine the time it takes for each robot to complete the task. This information will then be used to decide which robot to purchase.

   What type of test should the company use for this evaluation, and why?

   - Regression, since the time taken is scale data and we are therefore comparing scale vs scale data
   - <u>ANOVA, since the companies are nominal data, and the time taken is scale data</u>
   - T-test, since the companies are nominal data, and the time taken is scale data

2. The table below shows the results of the test performed, with each entry showing the time taken (in seconds) to perform the task during a particular trial.

Vendor A	Vendor B	Vendor C
8	8	10
10	9	9
9	9	10
11	8	11
10	10	9

   We will use this data to perform an ANOVA analysis. Based on the analysis, which vendor should the company select`?

   - <u>Vendor A</u>
   - Vendor B
   - Vendor C

Summary of Data						
	Treatments					
	1	2	3	4	5	Total
N	5	5	5			15
∑X	48	44	49			141
Mean	9.6	8.8	9.8			9.4
$\sum X^2$	466	390	483			1339
Std.Dev.	1.1402	0.8367	0.8367			0.9856

Result Details				
Source	*SS*	*df*	*MS*	
Between-treatments	2.8	2	1.4	*F* = 1.55556
Within-treatments	10.8	12	0.9	
Total	13.6	14		

The *F*-ratio value is 1.55556. The *p*-value is .250789. The result is *not* significant at $p < .05$.

3. In the above analysis, how many degrees of freedom are available for the R sub-table?

   11

   15 samples – 1 degree of freedom – 3 degrees of freedom for treatment averages – 1 degree of freedom for global average.

4. Suppose Vendor D also submits a robot consideration. The robot costs 1.2 Million CHF, and achieves time trial results of 7, 6, 8, 8, and 8 seconds. Under this new scenario, which robot should the company purchase?

- Vendor A
- Vendor B
- Vendor C
- Vendor D

Summary of Data

	Treatments					
	1	2	3	4	5	Total
N	5	5	5			20
∑X	48	44	49			178
Mean	9.6	8.8	9.8			8.9
$\sum X^2$	466	390	483			1616
Std.Dev.	1.1402	0.8367	0.8367			1.2937

Result Details

Source	*SS*	*df*	*MS*	
Between-treatments	17.8	3	5.9333	$F = 6.78095$
Within-treatments	14	16	0.875	
Total	31.8	19		

The *f*-ratio value is 6.78095. The *p*-value is .003673. The result is significant at $p < .05$.

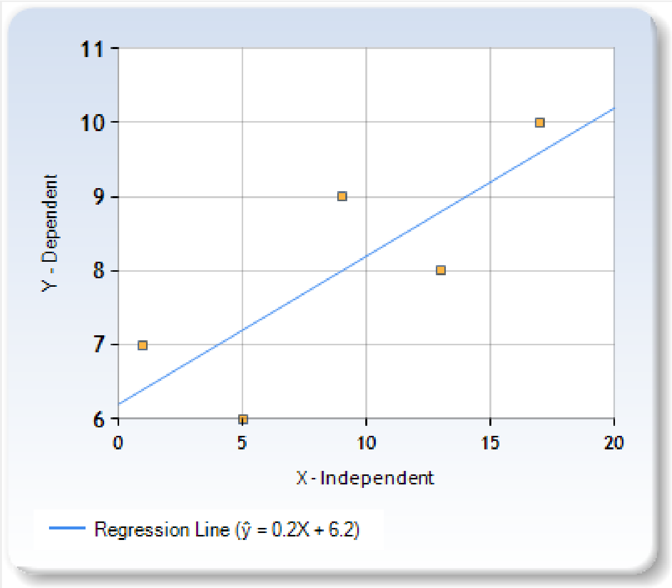
5. An economist in England performed analysis of household income vs household expenditures. She obtained the following table:

Income (in 1000 £)	Expenditure (in 1000 £)
1	7
5	6
9	9
13	8
17	10

Perform linear regression of this data to obtain a best-fit line. What is the slope of the line?

0.2

X Values	*Y* Values
1 5 9 13 17 M: 9	7 6 9 8 10 M: 8



Estimate	Estimated *Y*
7.75	7.75

Calculation Summary

Sum of X = 45
Sum of Y = 40
Mean X = 9
Mean Y = 8
Sum of squares (SS_X) = 160
Sum of products (SP) = 32

Regression Equation = $\hat{y} = bX + a$

$b = SP/SS_X = 32/160 = 0.2$

$a = M_Y - bM_X = 8 - (0.2*9) = 6.2$

$\hat{y} = 0.2X + 6.2$

6. What is the Pearson correlation coefficient for this line?

   0.8

Result Details & Calculation	Key
X Values $\sum = 45$ Mean = 9 $\sum(X - M_x)^2 = SS_x = 160$ *Y Values* $\sum = 40$ Mean = 8 $\sum(Y - M_y)^2 = SS_y = 10$ *X and Y Combined* $N = 5$ $\sum(X - M_x)(Y - M_y) = 32$ *R Calculation* $r = \sum((X - M_y)(Y - M_x)) / \sqrt{((SS_x)(SS_y))}$ $r = 32 / \sqrt{((160)(10))} = 0.8$ *Meta Numerics (cross-check)* r = 0.8	*X*: X Values *Y*: Y Values M_x: Mean of X Values M_y: Mean of Y Values $X - M_x$ & $Y - M_y$: Deviation scores $(X - M_x)^2$ & $(Y - M_y)^2$: Deviation Squared $(X - M_x)(Y - M_y)$: Product of Deviation Scores

The value of R is 0.8.

7. For a family to survive, the household income must at least equal to the household expenditure. Using the best fit line, determine the minimum household income for a family to survive, in thousands of £.

   7.75

   *Solve the equation val = 0.2*val + 6.2*
   *0.8*val = 6.2*
   Val = 6.2/0.8 = 7.75

8. The economist uses the data to predict the expected household consumption for a family earning 6000 £ and for a family earning 25000 £. Would you trust these estimates?

   - No, I would not trust either estimate since the R value is poor
   - Yes, I would trust both estimates since the R value is good
   - <u>I would only trust the value at 6000 £ since we shouldn't extrapolate too far beyond the bounds of the data</u>
   - I would only trust the value at 25000 £ since the 6000 £ data is close to an influential outlier.

9. A statistician performs an ANOVA analysis. He provides us with the following table. Unfortunately, some data in the table is missing. The missing data is indicated by a letter a-g in the table below.

	Sum of Squares	Degrees of Freedom	Mean Square	F
Between Treatments	2.124	a	0.708	0.75
Within Treatments	b	20	c	
Total	d	e		

What is the value of a?

3 (i.e., 2.124 / 0.708)

10. What is the value of b?

18.88 (answer from part 11, i.e., 0.944 * 20)

11. What is the value of c?

0.944 (i.e., 0.708/0.75)

12. What is the value of d?

21.004 (sum of 2.124 and b)

13. What is the value of e?

23 (a + 20)

14. In the silicon chips used in automative electronics, the wiring is commonly done using Aluminium with 0.85% silicon. The silicon content of each of 25 randomly selected wiring samples was determined, and the following results were obtained:

Silicon content:

Mean: 0.888%
Standard deviation: 0.1807%
T score: 1.05
P value: 0.3

What type of test was used to obtain these results?

- 1 sample t test
- 2 sample t test
- Paired t test
- z test

15. In the above test, what is the null hypothesis?

- $\mu = 0.85$
- $\mu \lt\gt 0.85$
- $\mu = 0.888$, $\sigma = 0.1807$

16. Is this analysis one-tailed or two tailed?

- One-tailed
- Two-tailed

17. Does the sample set fall within the expected specifications for the wiring material?

- Yes
- No

18. As you may know, the traditional method of pouring champagne is to hold the glass vertically, under the belief that this preserves tiny gas bubbles that improve flavor and aroma.   This is different from beer, where the class is usually tilted to reduce the frothing.

Vertical

Tilted

To determine the best way to pour, measurements were performed while pouring champagne at different temperatures, and the following data was obtained:

Temperature (°C)	Pour type	n	CO2 Loss (g/L)	SD
18	Vertical	4	4	0.5
18	Tilted	4	3.7	0.3
12	Vertical	4	3.3	0.2
12	Tilted	4	2	0.3

At 18°C, is there a difference between a vertical pour or a tilted pour?

- Yes
- No

Calculate by using

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} \qquad s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

= (4-3.7)/(sqrt(0.5^2/4+0.3^2/4)

T Score: 1.02

DF: 6

Significance Level:

- ○ .01
- ◉ .05
- ○ .10

One-tailed or two-tailed hypothesis?:

- ◉ One-tailed
- ○ Two-tailed

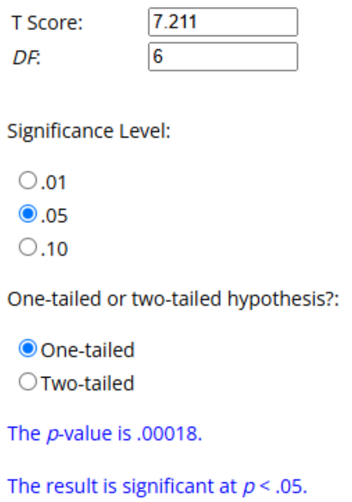
The *p*-value is .173541.

The result is *not* significant at *p* < .05.

19. Repeat this analysis at 12°C. What is the p-value?

0

=(3.3-2)/sqrt(0.2^2/4+0.3^2/4)



20. So, to retain bubbles with a chilled champagne, should you pour vertically or tilted?

- Vertically
- <u>Tilted</u>

21. During breastfeeding, women often temporarily lose bone mass, since the body uses calcium to produce milk. To test this hypothesis, data was collected from 10 women while they were breastfeeding, and again shortly after their children had grown up enough to stop breastfeeding. The data is provided below.

Female	Bone Mineral Content (g)	
	Lactating	Post-lactation
1	1928	2126
2	2549	2885
3	2825	2895
4	1924	1942
5	1628	1750
6	2175	2184
7	2114	2164
8	2621	2626
9	1843	2006
10	2541	2627

Perform both paired and 2-sample t-tests on this data to determine if there is in-fact mineral loss during lactation.

What is the t-value for the paired t-test?

3.21

Treatment 1	Treatment 2	*Diff* (T2 - T1)	*Dev* (Diff - M)	*Sq. Dev*
1928	2126	198	92.3	8519.29
2549	2885	336	230.3	53038.09
2825	2895	70	-35.7	1274.49
1924	1942	18	-87.7	7691.29
1628	1750	122	16.3	265.69
2175	2184	9	-96.7	9350.89
2114	2164	50	-55.7	3102.49
2621	2626	5	-100.7	10140.49
1843	2006	163	57.3	3283.29
2541	2627	86	-19.7	388.09
		M: 105.7		S: 97054.1

Significance Level:

- ○ 0.01
- ◉ 0.05
- ○ 0.10

One-tailed or two-tailed hypothesis?:

- ◉ One-tailed
- ○ Two-tailed

<u>Difference Scores Calculations</u>

Mean: 105.7

$\mu = 0$

$S^2 = SS/df = 97054.1/(10-1) = 10783.79$

$S^2_M = S^2/N = 10783.79/10 = 1078.38$

$S_M = \sqrt{S^2_M} = \sqrt{1078.38} = 32.84$

<u>T-value Calculation</u>

$t = (M - \mu)/S_M = (105.7 - 0)/32.84 = 3.22$

The value of *t* is 3.218765. The value of *p* is .00526. The result is significant at $p < .05$.

22. What is the t-value for the 2-sample t-test?

0.588

Significance Level:

- ○ .01
- ◉ .05
- ○ .10

One-tailed or two-tailed hypothesis?:

- ◉ One-tailed
- ○ Two-tailed

<u>Difference Scores Calculations</u>

Treatment 1

N_1: 10
$df_1 = N - 1 = 10 - 1 = 9$
M_1: 2214.8
SS_1: 1416571.6
$s^2_1 = SS_1/(N - 1) = 1416571.6/(10-1) = 157396.84$

Treatment 2

N_2: 10
$df_2 = N - 1 = 10 - 1 = 9$
M_2: 2320.5
SS_2: 1484580.5
$s^2_2 = SS_2/(N - 1) = 1484580.5/(10-1) = 164953.39$

<u>T-value Calculation</u>

$s^2_p = ((df_1/(df_1 + df_2)) * s^2_1) + ((df_2/(df_2 + df_2)) * s^2_2) = ((9/18) * 157396.84) + ((9/18) * 164953.39) = 161175.12$

$s^2_{M_1} = s^2_p/N_1 = 161175.12/10 = 16117.51$
$s^2_{M_2} = s^2_p/N_2 = 161175.12/10 = 16117.51$

$t = (M_1 - M_2)/\sqrt{(s^2_{M_1} + s^2_{M_2})} = -105.7/\sqrt{32235.02} = -0.59$

The *t*-value is -0.58872. The *p*-value is .281681. The result is *not* significant at $p < .05$.

23. Does lactation cause statistically significant bone mineral loss?

- <u>Yes</u>
- No

24. A sample of students was categorized according to gender and eye color, as listed in the table below:

Gender	Eye Color				
	Blue	Brown	Green	Hazel	TOTAL
Male	370	352	198	187	
Female	359	290	110	160	
TOTAL					

Suppose that one of these students is randomly selected. Let F denote the event that the selected individual is a female, and BL, BR, G, and H represent the events that he or she has blue, brown, green, and hazel eyes, respectively.

Calculate P(F)

0.4536, i.e., (359+290+110+160)/(sum of all)

25. Calculate $P(F \cap G)$

0.0543 (i.e, 110/2026)

26. Are events F and G independent?

- Yes
- No

A and *B* are independent if and only if (iff)

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.8)$$

*So, we check by seeing if P(F)*P(G)=0.0543, which it is not, since P(F)=0.4536 and P(G)=0.152, so P(F)*P(G)=0.068*

27. If the selected individual has green eyes, what is the probability that he or she is a female?

0.3571 (i.e., 110/308)

28. If the selected individual is female, what is the probability that she has green eyes?

0.1197 (i.e., 110/919)

29. A restaurant serves three fixed-price dinners costing CHF12, CHF15, and CHF20. For a randomly selected brother and sister dining at this restaurant, let X = the cost of the brother's dinner and Y = the cost of the sister's dinner. The joint pmf of X and Y is given in the following table:

p(x,y)		y		
		12	15	20
	12	0.05	0.05	0.1
x	15	0.05	0.1	0.35
	20	0	0.2	0.1

What is the probability that the brother's and sister's dinner cost at most CHF15 each?

0.25 (i.e, 0.05 +0.05 + 0.05 + 0.1)

30. Are X and Y independent?

- Yes
- <u>No</u>

$p_X(12) \cdot p_Y(12) = (.2)(.1) \neq .05 = p(12,12)$, so X and Y are not independent. (Almost any other (x, y) pair yields the same conclusion).

31. What is the expected total cost of the dinner for the two people?

33.35 i.e, $E(X+Y) = \Sigma\ \Sigma\ [(x + y)\ p(x, y)] = 33.35$.

= 0.05(12+12) + 0.05(15+12)+0.1(20+12)….(0.1(20+20)

32. Suppose that when a pair of diners opens fortune cookies at the conclusion of the meal, they find the message "You will receive as a refund the difference between the cost of the more expensive and the less expensive meal that you have chosen." How much would the restaurant expect to refund?

3.85 i.e.e, $E(X - Y) = \Sigma\ \Sigma$ abs$(x - y)\ p(x, y)] = 3.85$

= 0.05*3+0.05*3+0.2*5+0.1*8+0.35*5

33. Let $\tau > 0$ and let $X_{1,}$ $X_{2,\ldots}X_n$ be random samples from a distribution with a pdf given by:

$$f(x;\tau) = \frac{\tau^5}{8} x^{14} e^{-\tau x^3}, \qquad x > 0.$$

Which of the following is the maximum likelihood estimator of τ?

- $\hat{\tau} = \frac{n}{\sum_{i=1}^{n} X_i^3}$
- $\hat{\tau} = \frac{5n}{\sum_{i=1}^{n} X_i^3}$
- $\hat{\tau} = \frac{n}{5 \cdot \sum_{i=1}^{n} X_i^3}$
- $\hat{\tau} = \frac{\frac{5}{8}n}{\sum_{i=1}^{n} X_i^3}$

We write the log function as (only working here for 1 sample; obviously, we would actually do this for n samples)

Ln(f)= ln(tau^5)-ln(8) + ln(x^14) + ln(e^(-tau*x^3)

= 5ln(tau) – ln(8) + 14ln(x) – tau*x^3

To maximize, we can take the derivative wrt tau, which is:

5/tau -x^3, which we set to zero.

5/tau-x^3= 0

5/tau=x^3

Tau = 5/x^3

This calculation was done for only 1 sample. If we have n samples, we would obtain answer 2 above.

34. Let $X_1, X_2 \ldots X_n$ be a random sample from a population with pdf:

$$f(x|\theta) = \frac{1}{\theta} x^{(1-\theta)/\theta}$$

Where $0 < x < 1$, $0 < \theta < \infty$

Obtain a method of moments estimator for θ, and use this to calculate an estimate when x_1 = 0.1, x_2 = 0.22, x_3 = 0.54, x_4 = 0.36. What is the value of the estimator?

2.2787

Since there is one parameter, we just need to find the mean of the distribution and equate it to the sample mean.

The mean of the function is

Step 1: Find the method of moments estimator (MME)

The method of moments equates the **sample mean** to the **population mean**.

We compute the **expected value** of X:

Let's calculate $\mathbb{E}[X]$:

$$\mathbb{E}[X] = \int_0^1 x \cdot f(x \mid \theta)\, dx = \int_0^1 x \cdot \frac{1}{\theta} x^{\frac{1-\theta}{\theta}}\, dx = \frac{1}{\theta} \int_0^1 x^{1+\frac{1-\theta}{\theta}}\, dx$$

Simplify the exponent:

$$1 + \frac{1-\theta}{\theta} = \frac{\theta + 1 - \theta}{\theta} = \frac{1}{\theta}$$

So:

$$\mathbb{E}[X] = \frac{1}{\theta} \int_0^1 x^{1/\theta} dx = \frac{1}{\theta} \cdot \left[\frac{x^{1/\theta+1}}{1/\theta + 1} \right]_0^1 = \frac{1}{\theta} \cdot \frac{1}{1/\theta + 1}$$

Now simplify:

$$\mathbb{E}[X] = \frac{1}{\theta} \cdot \frac{1}{1/\theta + 1} = \frac{1}{\theta(1/\theta + 1)} = \frac{1}{1+\theta}$$

Step 2: Set the sample mean equal to population mean

Let the sample mean be:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Set:

$$\bar{X} = \mathbb{E}[X] = \frac{1}{1+\theta}$$

Solve for θ:

$$\bar{X} = \frac{1}{1+\theta} \Rightarrow 1+\theta = \frac{1}{\bar{X}} \Rightarrow \theta = \frac{1}{\bar{X}} - 1$$

Step 3: Plug in data to compute MME

Given:

$x_1 = 0.1, x_2 = 0.22, x_3 = 0.54, x_4 = 0.36$

Compute the sample mean:

$$\bar{X} = \frac{0.1 + 0.22 + 0.54 + 0.36}{4} = \frac{1.22}{4} = 0.305$$

Now plug into the estimator:

$$\hat{\theta}_{\text{MM}} = \frac{1}{0.305} - 1 \approx 3.2787 - 1 = \boxed{2.2787}$$