Midterm exam - practice

1. True or false question - only circle "true " or "false".
   (a) True or false? In a Graeco-Latin Square design two nuisance factors are blocked.

   (b) True or false? Type II error is the probability that we fail to reject a false null hypothesis.

   (c) True or false? From the central limit theorem, if a sample size is large, then the shape of a histogram of the sample will be approximately normal, even if the population distribution is not normal.

2. The diameter of a ball bearing was measured by 12 inspectors, each using two different kinds of calipers. The results were:

Inspector	Caliper 1	Caliper 2	Difference	Difference^2
1	0.265	0.264	0.001	0.000001
2	0.265	0.265	0.000	0
3	0.266	0.264	0.002	0.000004
4	0.267	0.266	0.001	0.000001
5	0.267	0.267	0.000	0
6	0.265	0.268	-0.003	0.000009
7	0.267	0.264	0.003	0.000009
8	0.267	0.265	0.002	0.000004
9	0.265	0.265	0.000	0
10	0.268	0.267	0.001	0.000001
11	0.268	0.268	0.000	0
12	0.265	0.269	-0.004	0.000016
			$\sum = 0.003$	$\sum = 0.000045$

Part of Output

Paired T-Test

Paired T for Caliper 1 - Caliper 2

	N	Mean	StDev	SE Mean
Caliper	12	0.266250	0.001215	0.000351
Caliper	12	0.266000	0.001758	0.000508
Difference	12	0.000250	0.002006	0.000579

(a) Construct a hypothesis test for the difference between the means of the population of measurements represented by the two samples.

$$H_0 : \mu_d = 0$$
$$H_1 : \mu_d \neq 0$$

(b) Find the t-value and *P*-value for the test in part (a).

$$t=\bar{d}/(S_d/\sqrt{n})= \bar{d}/\text{SE mean}$$
$$t=0.00025/0.000579=0.4318,\ \ df=11,\ p\text{-value}=0.6742$$

From t table we see that $0.4318<t_{0.75,11}=0.697$, which means p-value >2*(1-0.75)=0.5

(c) Construct a 95 percent confidence interval on the difference in the mean diameter measurements for the two types of calipers.

$$\bar{d} - t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}} \le \mu_D\left(=\mu_1-\mu_2\right) \le \bar{d} + t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}}$$
$$0.00025-2.201\frac{0.002}{\sqrt{12}} \le \mu_d \le 0.00025+2.201\frac{0.002}{\sqrt{12}}$$
$$-0.00102 \le \mu_d \le 0.00152$$

3. The following output was obtained from a computer program that performed ANOVA on an experiment

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	3	**?**	**?**	**?**	**?**
B	**?**	280.378	**?**	**?**	**?**
Error	9	58.797	**?**		
Total	15	347.653			

(a) Fill in the blanks in the ANOVA table. You can use bounds on the *P*-values.

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	3	**8.478**	**2.826**	**0.4326**	**0.7348**
B	**3**	280.378	**94.460**	**14.4589**	**0.0009**
Error	9	58.797	**6.533**		
Total	15	347.653			

(b) How many levels were used for factor *B*?

4 levels

(c) State the statistical model with constrains.

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, i = 1, \dots, 4, j = 1, \dots, 4$$

$$\sum_{i=1}^{4} \tau_i = 0, \qquad \sum_{j=1}^{4} \beta_j = 0, \quad \varepsilon_{ij}\ i.i.d. \sim N(0, \sigma^2)$$

(d) What conclusions would you draw about this experiment?

Factor *B* is statistically significant with a *P*-value of 0.009 while factor *A* is not significant.

4. An experiment was conducted to test the effects of nitrogen fertilizer on lettuce production. Five rates of ammonium nitrate were applied to four replicate plots in a design below. The data are the number of heads of lettuce harvested from the plot.

Treatment (lbs N/acre)	Heads of Lettuce/plot	Mean	St. D.
0	100, 120, 90, 140	112.50	22.1736
50	130, 135, 145, 175	146.25	20.1556
100	148, 140, 150, 158	149.00	7.3937
150	145, 160, 165, 162	158.00	8.9069
200	130, 150, 155, 160	148.75	13.1498

The following ANOVA table is obtained from SAS.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	?	4939.300000	?	?	?
Error	?	?	?		
Corrected Total	?	8553.800000			

(a) Identify what design this is.
Completely randomized design.

(b) State the statistical model with constrains.

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, i = 1, \dots, 4, j = 1, \dots, 4$$

$$\sum_{i=1}^{4} \tau_i = 0, \qquad \varepsilon_{ij}\ i.i.d. \sim N(0, \sigma^2)$$

(c) Fill in the missing values (indicated by "?") in the ANOVA table.

ANOVA table:

Source	DF	Sum of Squares	Mean Square	F-value	Pr > F
Model	4	4939.3	1234.8	5.1244	< 0.01
Error	15	3614.5	240.9667		
Total	19	8553.8			

(d) Test whether the nitrogen fertilizer affects the lettuce production at the 0.05 level of significance.

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

H_a : At least one μ_i is different.

P-value < 0.01

Reject H_0.

(e) What is the estimate of the population variance?

$$\hat{\sigma}^2 = 240.9667$$