Name: ______________________________ ID #: ______________________________

CENG 114 – Probability and Statistical Methods for Engineers

Professor Shyue Ping Ong

Final Exam

Winter, Mar 20 2018

Instructions:

1. Write your name on the top right of this page.
2. Please answer **ALL** questions as far as possible.
3. You have a total of 180 mins exactly.
4. Write your answers in the space provided following the questions. If you need additional space, you may request for additional blank sheets of paper during the exam. If you use additional blank sheets, you should mark the top left of each sheet with your name, ID#, question number, and today's date, e.g., "John Smith, A1234567, Q1, Feb 27 2018".
5. Indicative maximum points for each sub-question are provided only as a reference. Actual grading points may vary.
6. Write neatly and legibly. Illegible answers will be considered as wrong.

Question 1 (20 points): ____________

Question 2 (20 points): ____________

Question 3 (20 points): ____________

Question 4 (20 points): ____________

Question 5 (20 points): ____________

This exam has a total of 10 pages, including the cover page.

Qn 1.

A study on the protopectin content of tomatoes during storage was carried out. Four storage times were selected, and samples from four lots of tomatoes were analyzed. The protopectin content (expressed as hydrochloric acid soluble fraction mg/kg) is in the table below.

	Lot				
Storage time	**1**	**2**	**3**	**4**	
0 days	1694	1123	1106	1116	**5039**
7 days	1000	818	407	462	**2687**
14 days	650	422	420	409	**1901**
21 days	415	394	356	351	**1516**
	3759	**2757**	**2289**	**2338**	**11143**

i. (10 points) Carry out an ANOVA test at $\alpha = 0.05$ to test whether there is a significant difference in protopectin content with different storage times. State all your hypotheses clearly, and complete the following ANOVA table.

	SS	**dof**	**MS**	**F**
Treatment				
Within				
Subject				
Error				
Total				

ii. (6 points) Perform a Tukey's HSD test to ascertain which storage times resulted in significantly different protopectin content.

Solutions:

i.

$H_0: u_0 = u_7 = u_{14} = u_{21}$

$H_1: H_0 \; is \; false$

$$\sum x^2 = 10169337, \quad \sum \frac{T^2}{n} = 9630886.75, \frac{G^2}{N} = 7760403.063, \sum \frac{{T_{subject}}^2}{I} = 8109223.75$$

	SS	dof	MS	F
Treatment	1870483.688	3	623494.563	29.59
Within	538450.25	12		
Subject	348820.688	3		
Error	189629.563	9	21069.951	
Total	2408933.938	15		

$dof_{Tr} = 3, dof_E = 9, F_{crit} = 3.863, \; F = 29.59, F > F_{crit}, so \; reject \; H_0$

ii.

$$I = 4, dof_E = 9, q = 4.415, HSD = q\sqrt{\frac{MSE}{n}} = 4.415 * \sqrt{\frac{21069.95}{4}} \rightarrow HSD = 320.43$$

$u_0 = 1259.75, u_7 = 671.75, u_{14} = 475.25, u_{21} = 379.0$

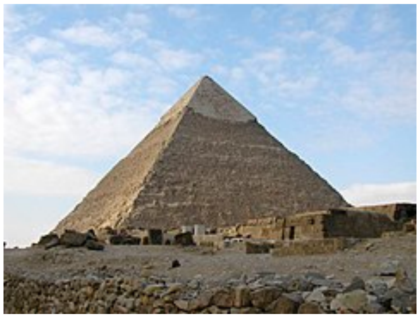
$|\,u_0 - u_7| > HSD, |\,u_0 - u_{14}| > HSD, |\,u_0 - u_{21}| > HSD$

$|\,u_7 - u_{14}| < HSD, |\,u_7 - u_{21}| < HSD, |\,u_{14} - u_{21}| < HSD$

So the following storage time *pairs* result in significantly different protopectin content:

$|\,u_0 - u_7| > HSD, |\,u_0 - u_{14}| > HSD, |\,u_0 - u_{21}| > HSD$

Qn 2.



Dr Ross Geller, a famous archaeologist, is attempting to estimate the volume of the Great Pyramid of Giza. Using his instruments, he determines that the dimensions are $h = 138 \pm 2$ m, $l = 230 \pm 1$ m, $w = 232 \pm 1$ m, where h, l and w refers to the height, length and width respectively. The volume of the pyramid is given by:

$$V = \frac{l \times w \times h}{3}$$

i. (6 points) What is the estimated volume and the uncertainty?

ii. (9 points) Instead of assuming that the lengths and width are independent, Dr Geller decides to assume that this is a square pyramid. He then takes the average of his measurements of l and w and uses that as the side of the square base. What is the estimated volume and the uncertainty in this case?

iii. (5 points) To improve his estimate of the height, Dr Gellar decides to use some trignometry. He stands 500 m from the center of the pyramid and was able to determine that the angle between the tip of the pyramid and the ground is 0.269 ± 0.001 radians. Assuming the same approach to measuring the volume as part (i), what is the uncertainty in the estimated volume? (Hint: The derivative of tan(x) is $\sec^2(x)$.)

Solutions:

i.

$$V_{estimate} = 230 * 232 * \frac{138}{3} = 2454560$$

$$\sigma_V = \sqrt{(\frac{\partial V}{\partial l})^2 * \sigma_l{}^2 + (\frac{\partial V}{\partial w})^2 * \sigma_w{}^2 + (\frac{\partial V}{\partial h})^2 * \sigma_h{}^2}$$

$$\sigma_V = \sqrt{(\frac{232 * 138}{3})^2 * 1^2 + (\frac{230 * 138}{3})^2 * 1^2 + (\frac{230 * 232}{3})^2 * 2^2} = 38617.225$$

ii.

$$s = \frac{l + w}{2} = 231 \pm 1, V = \frac{s^2 * h}{3} = \frac{(l + w)^2 * h}{12}$$

$$V_{estimate} = \frac{231^2 * 138}{3} = 2454606$$

$$\sigma_V = \sqrt{(\frac{\partial V}{\partial l})^2 * \sigma_l{}^2 + (\frac{\partial V}{\partial w})^2 * \sigma_w{}^2 + (\frac{\partial V}{\partial h})^2 * \sigma_h{}^2}$$

$$\sigma_V = \frac{1}{12}\sqrt{(2(l+w)h)^2 * \sigma_l{}^2 + (2(l+w)h)^2 * \sigma_w{}^2 + ((l+w)^2)^2 * \sigma_h{}^2}$$

$$\sigma_V = \frac{1}{12}\sqrt{(2 * 462 * 138)^2 * 1 + (2 * 462 * 138)^2 * 1 + 462^4 * 4} = 38617.784$$

iii.

$$h = 500 * \tan\theta\,, V = \frac{500 * l * w * tan\theta}{3}$$

$$\sigma_V = \sqrt{(\frac{\partial V}{\partial l})^2 * \sigma_l{}^2 + (\frac{\partial V}{\partial w})^2 * \sigma_w{}^2 + (\frac{\partial V}{\partial \theta})^2 * \sigma_\theta{}^2}$$

$$\sigma_V = \frac{500}{3} * \sqrt{(63.958)^2 * 1^2 + (63.407)^2 * 1^2 + (55350.564)^2 * 0.001^2} = 17582.22$$

Qn 3.

The Hall-Petch equation describes the relationship between the yield strength of a metal and the grain size.

$$\sigma_y = \sigma_0 + \frac{k}{\sqrt{d}}$$

where σ_y is the yield stress, σ_0 is a materials constant for the starting stress for dislocation movement, k is the strengthening coefficient, and d is the average grain diameter. A metallurgist is investigating this relationship in Ni_3Al alloys and collected the following data:

σ_y**(MPa)**	***d*** **(μm)**					
999	6					
839	10					
615	30					
550	48					

508	70					
474	100					
445	150					
436	170					

i. (6 points) By performing a linear least squares regression, determine σ_0 and k for Ni_3Al to to 3 significant figures and specifying the units.

ii. (4 points) We may assume that the grains in a sample of Ni_3Al are normally distributed. After performing 30 careful measurements, a student determined that 50% of the grains have a size of less than 200 μm, while 74.86% of the grains have a size of less than 220 μm. What is the mean and standard deviation of the grain size?

iii. (4 points) Construct a 99% confidence interval for the mean grain size based on the measurements in part (ii).

iv. (2 points) What is the estimated yield strength of the Ni_3Al sample in (ii)?

v. (4 points) In a manufacturing process, Ni_3Al alloy samples are made with average grain sizes uniformly distributed from 100-200 μm. Calculate the probability that the yield strength of a randomly chosen sample is less than 450 MPa.

Solutions:

i.

$$\sigma_y = kd^{-0.5} + \sigma_0 \; or \; Y = kX + \sigma_0$$

$$k = \frac{SP}{SS_X} = 1694.369 \; (MPa \cdot um^{0.5})$$

$$\sigma_0 = \bar{Y} - k\bar{X} = 305.539 \; (MPa)$$

ii.

$$P(X < 200) = 0.5 \rightarrow Z = \frac{200-u}{\sigma} = 0 \rightarrow \; u = 200 \text{ (um)}$$

$$P(X < 220) = 0.7486 \rightarrow Z = \frac{220-u}{\sigma} = 0.67 \rightarrow \sigma = 29.850 \text{ (um)}$$

iii.

$$200 \pm 2.58\frac{29.85}{\sqrt{30}} = (185.94, 214.06) \quad \text{(um)}$$

iv.

$$\sigma_y = kd^{-0.5} + \sigma_0 = 305.539 + \frac{1694.369}{\sqrt{200}} = 425.349 \text{ (MPa)}$$

v.

$$\sigma_y = \frac{1694.369}{\sqrt{d}} = 450 \rightarrow d = 137.567 \text{ (MPa)}$$

$$P = \frac{137.567 - 100}{200 - 100} = 37.567\%$$

Qn 4.

Americium (^{241}Am, 241 being the atomic mass) is the only synthetic element that is widely found in households. The most common type of smoke detector uses ^{241}Am in the form of AmO_2 as its source of ionizing radiation. A typical smoke detector contains about 0.29 milligrams of ^{241}Am. ^{241}Am has a half-life of 432.2 years, and decays by emitting an α particle into ^{237}Np, which has a much longer half-life of 2.14 million years.

$$^{241}\text{Am} \rightarrow {}^{237}\text{Np} + \alpha$$

i. (2 points) What is the rate constant of the radioactive decay of ^{241}Am?

ii. (4 points) Calculate the ratio of ^{237}Np to ^{241}Am after 30 years.

iii. (8 points) Estimate the number of α particles emitted per day by a typical smoke detector at the beginning of its life. (Avogadro's constant = 6.022×10^{23} mol^{-1})

iv. (6 points) A smoke detector goes off when there is light, medium and heavy smoke with probability of 0.1, 0.5 and 0.9 respectively. In a kitchen, the probability of light, medium and heavy smoke being generated (depending on what is being cooked) is 0.2, 0.7 and 0.1 respectively. The smoke detector went off on a randomly chosen day. What is the probably that it is due to light smoke being generated?

Solutions:

i.

$$t_{\frac{1}{2}} = 432.2 = \frac{ln2}{\lambda} \rightarrow \lambda = 0.0016\ (year^{-1})$$

ii.

$$t = 30, \frac{A_0(1 - e^{-\lambda t})}{A_0 e^{-\lambda t}} = 0.0493$$

iii.

t = 1/365, $A_0\left(1 - e^{-\lambda t}\right) = \frac{0.29*10^{-3}}{241} * \left(1 - e^{-\lambda t}\right) * 6.022 * 10^{23} = 3182909764244.3$

iv.

$$P = \frac{0.1 * 0.2}{0.1 * 0.2 + 0.5 * 0.7 + 0.9 * 0.1} = 0.043$$

Qn 5.

A cereal manufacturer claims that the sodium content per 300-gram box in a particular brand of organic cornflakes is less than 130 mg. A researcher decides to test these claims by carrying out measurements of the sodium content of twelve 300-gram boxes of organic cornflakes. The data (in mg) are as follows:

131.15, 130.69, 130.91, 129.54, 129.64, 128.77, 130.72, 128.33, 128.24, 129.65, 130.14, 129.29

i. (6 points) Perform a hypothesis test at $\alpha = 0.05$ to test the manufacturer's claim. State clearly your null and alternative hypotheses, the test statistic that you are using and your conclusions.

ii. (6 points) It turns out that the true sodium content in this brand of organic cornflakes have distribution $N(132, 9)$. Determine the probability of a type II error.

iii. (4 points) The manufacturer has a quality control system in place where 10 boxes are randomly selected out of every batch of 1000 for testing. The whole batch is rejected if more than 2 boxes exceed a sodium content of 135 mg. Assuming the distribution in part (ii), what is the probability that a batch will be rejected?

iv. (4 points) Estimate the probability that there are more than 170 boxes with a sodium content exceeding 135 mg in a batch of 1000.

Solutions:

i.

$H_0: u \leq 130$

$H_1: u > 130$

$$\because small\ sample\ size, t = \frac{\bar{X} - u}{\frac{s}{\sqrt{n}}} = \frac{129.756 - 130}{\frac{0.992}{\sqrt{12}}} = -0.853$$

$dof = 12 - 1 = 11$

From Student's t table, $\alpha = 0.05, one\ tailed, t_{crit} = +1.796, t < t_{crit}. So\ do\ not\ reject\ H_0$

ii.

$$\frac{\bar{X} - 130}{0.992/\sqrt{12}} = 1.796 \rightarrow \bar{X} = 130.51$$

$$P(X < 130.51) = P\left(Z < \frac{130.51 - 132}{\frac{3}{\sqrt{12}}}\right) = P(Z < -1.73) = 1 - 0.9582 = 0.0418$$

iii.

$$P(X > 135) = P\left(Z > \frac{135 - 132}{3}\right) = P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587$$

$$P = 1 - [C_0^{10} * 0.8413^{10} + C_1^{10} * 0.8413^9 * 0.1587 + C_2^{10} * 0.8413^8 * 0.1587^2] = 0.203$$

iv.

$$X \sim Bin(1000, 0.1587) \sim N(158.7, 133.51)$$

$$P(X > 170) = P\left(Z > \frac{170 - 0.5 - 158.7}{\sqrt{133.51}}\right) = P(Z > 0.93) = 1 - P(Z < 0.93)$$

$$= 1 - 0.8238 = 0.1762$$