BIOE: Biostatistics Course Fall 2017

Assignment 5

Due 16th November

Provide any code you used with the assignment.

1.[4] Write down the short-hand for a normal distribution with:

(a) mean 5 and standard deviation 3
(b) mean -100 and variance 10

$N(5, 25)$
$N(-100, 10)$

2.[10] A new treatment for pain relief of kidney gall stones is being tested on 50 patients. Traditional treatments relieve pain on average for 3.5 hours. On trials with the 50 patients, the new drug is found to result in 3.9 hours of pain relief with a standard deviation of 2.2 hours.

Determine at the 5% significance level whether the new treatment relieves pain for longer.

p-value =- 0.099
Test statistic = 1.28
H_o is not rejected, no difference.

3.[10] Two groups of lung cancer patients are treated. One group is given no treatment, the second group is given a new but experimental immunotherapy. The following data were collected after 4 months of treatment, the numbers indicate the relative volume of the cancer, i.e how big the tumor is, smaller is obviously better.

Group 1: (No treatment):	91.5	94.18	92.18	95.39	91.79	89.07	94.72	89.21
Group 2: (With treatment):	89.19	90.95	90.46	93.21	97.19	97.04	91.07	92.75

Assume the population variances are equal. Is there any difference in the mean outcomes between the two groups? Use alpha = 0.05 (Note $\alpha/2$ will be in each tail)

Doesn't make much difference whether they do equal or unequal variances:
Unequal Variances

DF : 13 (rounded down from 13.0278)
Test Statistic t = -0.354
Population 1 Population 2: P-Value = 0.729

Equal Variances

Pooled Standard Deviation: 2.7009
Pooled DF: 14

Test Statistic t= -0.35395
Population 1 Population 2: P-Value = 0.7294

4.[20] You might want to use software to assist you in following exercise.

a) Generate a sample of 40 random numbers where each random number has range from 0 to 9. Make sure you understand clearly how to use `randint` for this.
b) Test the null hypothesis at $\alpha = 0.1$ that the population of random numbers has a mean $\mu = 4.5$
c) Repeat a) and b) 10 times.
d) What is the critical value for rejecting H_o at $\alpha = 0.1$?
e) Create a table of your results with the following columns:

Sample Number	Sample Mean	z-score	Accept or Reject H_o

f) How many reject and accept conclusions do you see in the last column?
g) What pattern did you expect to see in the last column?

h) To get a better feel for this, repeat a) and b) **100** times and compute the average number of times you see a reject H_0
i) Is this in line with what you'd expect? If so why?

Critical value at $\alpha = 0.1 is 1.645$
They should see about 10% rejects

5.[10] A range of five proteins are measured using a very precise mass-spectroscopy technique which allows individual proteins to be counted. Proteins are counted in

two cell types, a stem cell and a differentiated heart cell. The differentiated heart cells are derived from the stem cells by stimulating differentiation. Determine whether the distribution of proteins in the heart cells has changed significantly from the same group of proteins in the stem cells. The entries in the table below give the number of proteins per cell in the stem and differentiated cells respectively.

Protein Type	P53	NC34	16SP	SOX	NANOG
Stem Cell	26	120	8	35	2
Differentiated Cell	10	95	16	9	3

	P53	**NC34**	**16SP**	**SOX**	**NANOG**	
Stem Cell	26 *21.22* (1.08)	120 *126.74* (0.36)	8 *14.15* (2.67)	35 *25.94* (3.17)	2 *2.95* (0.30)	**191**
Diff Cell	10 *14.78* (1.54)	95 *88.26* (0.52)	16 *9.85* (3.84)	9 *18.06* (4.55)	3 *2.05* (0.44)	**133**
	36	**215**	**24**	**44**	**5**	**324**

χ^2 = 18.457, df = 4, χ^2/df = 4.61, P(χ^2 > 18.457) = 0.0010

warning: some observed or expected frequencies are less than 5; thus the Central Limit Theorem may not apply and the resultant χ^2 may b

expected values are displayed in *italics*
individual χ^2values are displayed in (parentheses)

Determine whether the differentiated cells have a significantly different protein composition or not?

6.[16] In the following table, taken from the class notes, determine if there is a relationship between diet and cancer and heart disease?

	Cancer	Fatal Heart Dis-ease	Non-fatal Heart Disease	Healthy	Total
Healthy Diet	15	24	25	239	303
Mediterranean	7	14	8	273	302
Asian	4	3	16	260	283
Burger Diet	21	39	45	190	295
Total	47	80	94	962	1183

	Cancer	**Fat Heart**	**Non-Fatal Heart**	**Total**	
Healthy	15 *7.11* (8.77)	24 *12.10* (11.72)	25 *21.62* (0.53)	239 *262.18* (2.05)	**303**
Med	7 *7.08* (0.00)	14 *12.06* (0.31)	8 *21.55* (8.52)	273 *261.31* (0.52)	**302**
Asian	4 *6.64* (1.05)	3 *11.30* (6.09)	16 *20.19* (0.87)	260 *244.87* (0.93)	**283**
USA	21 *26.17* (1.02)	39 *44.55* (0.69)	94 *79.63* (2.59)	962 *965.64* (0.01)	**1116**
	47	**80**	**143**	**1734**	**2004**

χ^2 = 45.685, df = 9, χ^2/df = 5.08, P(χ^2 > 45.685) = 0.0000

warning: some observed or expected frequencies are less than 5; thus the Central Limit Theorem may not apply and the re

expected values are displayed in *italics*
individual χ^2values are displayed in (parentheses)