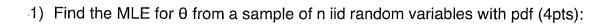
Student's Name:	SFU id:

Unless otherwise specified consider the 5% significance level.

Calculators are not permitted

You are permitted one page of double sided notes.

Page	Value	Score
	4	
3	6	
4	7	
5	6	
6	8	
7	6	
8	8	
9	9	·
IO	2	
Total	56	



$$f(y; heta) = rac{3}{ heta^3} y^2$$
 Under the constraint that for all y:  $0 < y < heta$ 

- 2. A researcher is attempting to measure out a triangular region of the sea floor for careful analysis. The triangle has 2 identical length sides of length  $\mu$  and identical 45° angles. The area of the triangle is then .5 ( $\mu^2$ ). The researcher is attempting to estimate the area of the triangle based on n independent measurements X1,X2,...Xn of  $\mu$ . Assume that each X<sub>i</sub> has mean  $\mu$  so that the measurement device is unbiased and the measurement is subject to measurement error variance  $\sigma^2$ .
- a) Show that  $(\frac{1}{2})$   $\bar{\chi}^2$  is not an unbiased estimator of the area of the triangle. (3pts)
- b) For what value of k is  $\overline{X}^2 kS^2$  an unbiased estimator of  $\mu^2$ ? (3pts)

Where we use:

$$S^{2} = \frac{1}{n-1} \left[ \sum X_{i}^{2} - \frac{(\sum X_{i})^{2}}{n} \right]$$

3. Given two independent samples of iid Normal random variates with different means and variances:  $y_1,...,y_n$  and  $x_1,...,x_m$ . If n=5 and m=7, derive an upper 90% confidence bound for the ratio  $\sigma_X^2/\sigma_Y^2$  (4 points for the derivation + 3 for the interval upper bound if  $S^2_x = 2$  and  $S^2_y = 4$ )

Hint 
$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$$
 has an F distribution.

(6points) Given a sample of 25 iid normal random variables with sample mean of 10 and sample standard deviation of 5 find a 90% CI for the mean. Find a 95% CI for the variance

A professor is trying to decide if one of two potential versions of an exam is more difficult than the other. The professor chooses 2 random samples of students and randomly assigns them to write either version 1 or version 2. The summaries from the samples are given below:

	Version 1	Version 2
n	50	50
mean	80	. 75
standard deviation	-10	10

a) What are the hypotheses to test? (2pts)

b) What do you conclude? Justify your answer with a p-value. (6 pts)

We are interested in examining the change in size of odd year pink salmon from the department of fisheries and oceans fishing records. Plots of the data (not shown here) suggest that a linear model would be appropriate to describe the trend between salmon weight and year. The 101 observations have the following summaries:

$$S_{xy} = -10000$$
  
 $S_{xx} = 200$   
mean of  $X = 100$   
mean of  $Y = 100$ 

a) What is the regression equation? (4 points)

- b) By how much does the average size of salmon seem to change in a 10 year span? (1 point)
- c) How large are the salmon predicted to be in the year 0? (1point)

The probability of being successfully winning a certain game of random chance is thought to be 40%. Last year there was a 25% win rate out of 2400 games. If these were observed from a random sample of equally likely and independent events, test the hypothesis that the actual odds of winning are less than the intended standard. (8 points)

The corrosion level of pipes is being tested after treatment with 4 different corrosion settings buried in 3 different soil types. Below is the output from R

```
> summary(lm(data ~ soil + coating))
lm(formula = data ~ soil + coating)
Residuals:
           10 Median
   Min
                          30
                                Max
-4.083 -2.417 -1.167 3.083 6.167
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)
              57.833
                          3.204 18.052 1.86e-06 ***
              -6.750
soil2
                          3.204
                                  -2.107
                                           0.0797
soil3
              -3.750
                           3.204
                                  -1.171
                                           0.2862
                          3.699
coating2
                                  -0.991
              -3.667
                                           0.3599
coating3
              -7.000
                           3.699
                                  -1.892
                                           0.1073
coating4
              -5.667
                          3.699
                                  -1.532
                                           0.1765
Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
Residual standard error: 4.531 on 6 degrees of freedom
Multiple R-squared: 0.587,
  Adjusted R-squared: 0.2429
F-statistic: 1.706 on 5 and 6 DF, p-value: 0.2664
```

> anova(lm(data ~ soil + coating))
Analysis of Variance Table

Response: data

Df Sum Sq Mean Sq F value Pr(>F) soil 91.500 45.750 2.2287 0.1889 coating 83.583 27.861 1.3572 0.3422 Residuals 6 123.167 20.528

- a) What is the expected value of corrosion for coating type 1 buried in soil type 2? (2pts)
- b) What is the interpretation of the value of the "(Intercept)" 57.833 in the R output? (1point)
- c) How many degrees of freedom are there for the soil treatment and how many for coating treatment (2points)?
- d) How many pipe pieces were tested in total across all treatments? (1point)
- e) If there a difference between mean corrosion levels across treatments? (3 points, must justify your conclusion)

Define p-value (2points)