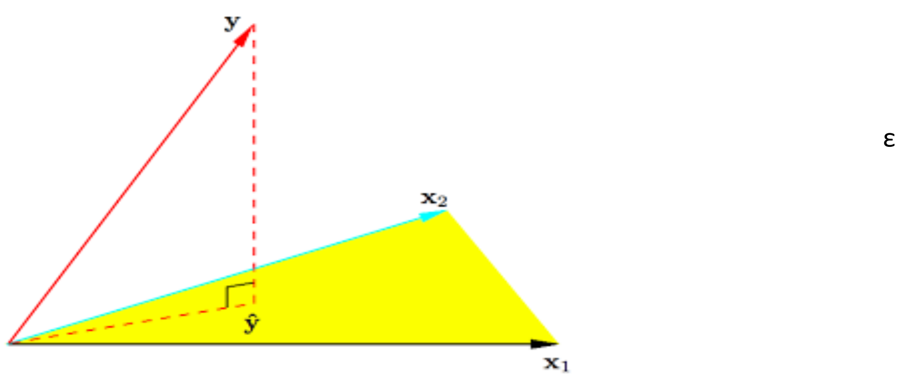
Applied Statistics Qualifier August 26, 2016 **Name___________________________**

STA-682 Portion (4 problems)

Assume 95% levels of confidence (α = .05) unless otherwise indicated. You should hand in your computer software output along with your analysis of each problem.

1. **Geometrically speaking, a Least Squares model may be depicted as follows:**



In which $\hat{y}$ is the projection of y into the column space of X.

Based on this information, prove the following:

A. **$\hat{y}$ and ε are independent (orthogonal).**
B. **The triangle equality: $\hat{y}^2 + \varepsilon^2 = y^2$**
C. **The projection of $X_2 \equiv X$ is invariant (nothing changes)**

2. **Consider a simple linear regression model**

$$Y = XB + \varepsilon$$

In which X is an nx2 Design Matrix, full column rank (consisting of the intercept and X_1), and $\varepsilon \sim N(0, \sigma^2 I)$ random variables.

In consideration of the following transformations, derive the revisions to the estimates of B_0 and B_1. In other words, how will B_0 and B_1 changed after these transformations are made?

A. **At the same time, multiply X_1 by 10 and divide Y by 10.**
B. **Inter-change X and Y (X_1 is now Y and Y is now X_1)**

3. **A 100 ft. rope is cut into 4 pieces of 25 ft. The length of each piece is measured 3 times by a different measurement device. The results are as follows:**

Measurement Device	Piece 1	Piece 2	Piece 3	Piece 4
1	24	24	25	27
2	23	27	22	24
3	25	22	26	27

Given that the sum of the 4 pieces must total 100 feet, which is a model constriction, use Least Squares in matrix format to do the following.

a. **State the Response Matrix, Design Matrix, Parameter Matrix, and any other matrices needed to create this estimation problem as a linear model.**

b. **Estimate the measurement of each of the 3 pieces using this model, subject to the constrictions as previously specified.**

c. **We wish to test if the measurement devices are accurate. Propose a test to determine if Ho: $\upsilon_1 = \upsilon_2 = \upsilon_3 = \upsilon_4$ versus H_1: At least one mean is not equal where υ_i represents the mean of Piece 1, 2, 3, 4 respectively.**

4. **One of the interesting properties of the Poissen distribution is that the variance is equal to the mean of the distribution. In other words, as the mean increases so does the variance. If the residuals were to follow a Poissen distribution, you would see the residual vs. fit plot exhibit a "fan" pattern. Obviously, this is a violation of a**

least squares requirement. Rather than considering a Box-Cox power transformation on Y ($Y^* = Y^\lambda$) to stabilize the residuals, propose a weighting structure for matrix (W) such that B= $(X'WX)^{-1}$ (X'W Y) to ensure homoscedasticity (constant variance)

For illustration purposes, assume that $\widehat{Y}$ = (1,2,3,4,……n)'

THE END