Stat 407 Quiz 2 Name _____

1. *(7 pts)* Answer these questions for the following variance-covariance matrix.

$$\mathbf{S} = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 3 & 3 \\ 4 & 3 & 20 \end{bmatrix}$$

   (a) *(1pt)* What is the dimension of the data? _____

   (b) *(1pts)* Which of the following two results is most likely those from PCA on the variance-covariance matrix, **S**? Explain your answer.

   _____

   _____

A

	PC1	PC2	PC3
V1	-0.74	0.01	0.67
V2	0.49	0.69	0.53
V3	-0.46	0.72	-0.52
Var	1.76	1.18	0.05

B

	PC1	PC2	PC3
V1	-0.32	0.69	-0.65
V2	-0.03	-0.69	-0.72
V3	-0.95	-0.21	0.24
Var	20.21	4.85	0.23

   (c) *(1pt)* Calculate the proportion of total variance explained by the first principal component of analysis A. _____

   (d) *(1pt)* What is the value of the second eigenvalue in analysis A? _____

   (e) *(1pt)* What is the value of the coefficient of V2 in the second eigenvector in analysis A? _____

   (f) *(1pts)* Draw the scree plot for results from analysis A. Be sure to label your axes.

   (g) *(1pts)* If an observation has the values [3 2 0] calculate its score on the first principal component of analysis A. (Assume that the means for each variables are 0.) _____

2. *(3 pts)* True or False.

   (a) When PCA is done on the correlation matrix, the total variance is equal to the number of variables. **T** or **F**

   (b) PCA summarizes only linear dependence. **T** or **F**

   (c) The first eigenvector describes the direction of maximum variance of the data. **T** or **F**